LIBRARY OF THE UNIVERSITY OF ALBERTA

For Reference

NOT TO BE TAKEN FROM THIS ROOM

Library of the University of Alberta, Edmonton, Alberta

Chalmers, John W.

A study of the effective and recognized vocabularies of Alberta students in grades VII to XII. 1935.

Ex libris
universitatis
albertaensis

A STUDY OF THE EFFECTIVE AND RECOGNIZED VOCABULARIES OF ALBERTA STUDENTS IN GRADES VII TO XII

BY

JOHN WEST CHALMERS.

A DISSERTATION

SUBMITTED TO THE COMMITTEE OF GRADUATE STUDIES, UNIVERSITY OF ALBERTA, IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF THE DEGREE OF MASTER OF ARTS.

EDMONTON, ALBERTA APRIL, 1935.

Thesis
1935
#30

TABLE OF CONTENTS

CHAPTER		PAGE
I	Purpose of Investigation	1
II	Sources of Data	2
III	Scoring the Themes	6
IV	Analyzing the written Vocabulary	18
V	Analyzing the Recognized Vocabulary	35
	Distribution by Grades	45
	Distribution by Ages	61
VI	Relationship between Vocabulary and Occupational Status	68
VII	Effect of Bilingualism on Vocabulary	76
VIII	Relationship between Intelligence and Vocabulary	80
IX	Sex Differences in Vocabulary	95
X	Vocabulary Scores and Other Achievements in English	97
XI	Conclusion	103

Digitized by the Internet Archive
in 2017 with funding from
University of Alberta Libraries

https://archive.org/details/chalmers1935

TABLES

		Page
I	Distribution of Subjects	4
II	Distribution of Hudelson Scores	8
III	Distribution of Written Vocabulary according to Comparative Frequency of Words	21
IV	Word Distribution according to Grade and frequency	33
V	Distribution of Vocabulary Scores by Grades	46
VIa	Overlapping of Vocabulary Scores by Grades	59a
VIb	Retardation and Acceleration of Vocabulary Scores by Grades	59b
VII	Distribution of Vocabulary Scores by Ages	62
VIII	Distribution of Vocabulary Scores according to Occupational Status	71
IX	Otis Interpretation Chart	82
Xa	Distribution of Intelligence Quotients by Grades	90
Xb	Distribution of Intelligence Quotients by Ages	90

DIAGRAMS

		Page
I	Distribution of Hudelson Scores in Grade VII	9
II	Distribution of Hudelson Scores in Grade VIII	10
III	Distribution of Hudelson Scores in Grade IX	11
IV	Distribution of Hudelson Scores in Grade X	12
V	Distribution of Hudelson Scores in Grade XI	13
VI	Distribution of Hudelson Scores in Grade XII	14
VII	Distribution of Words in Grade VII Compositions	24
VIII	Comparative distribution of Words in Written Compositions	25
IX	" " " " " (Continued)	26
X	Progressive changes in Relative Word Frequencies	27
XI	Distribution of Vocabulary Scores in Grade VII	47
XII	Distribution of Vocabulary Scores in Grade VIII	48
XIII	Distribution of Vocabulary Scores in Grade IX	49
XIV	Distribution of Vocabulary Scores in Grade X	50
XV	Distribution of Vocabulary Scores in Grade XI	51
XVI	Distribution of Vocabulary Scores in Grade XII	52
XVII	Grade Means for Vocabulary Test	55
XVIII	Grade Quartiles for Vocabulary Test	56
XIX	Age Means for Vocabulary Test	63
XX	Regression between Intelligence and Recognized Vocabulary	85
XXI	Distribution of Otis Scores in Grades VII to XII	86
XXII	Distribution of Vocabulary Scores in Grades VII to XII	87

CHAPTER I.

PURPOSE OF INVESTIGATION.

The purpose of this investigation is to study the vocabularies of high school students in the schools of the Province of Alberta, in an effort to find out the following:

I. To what extent do students use in writing the words they already know, i.e. what is the correspondence between vocabulary of expression and that of recognition?

II. How rapidly do students acquire new words and to what extent do they use those words?

III. What is the correlation between extent of vocabulary and native intelligence?

IV. To what extent does knowledge of another vernacular besides English affect command over the English language?

V. What, if any, is the effect of occupational status of a child's father upon the richness of the child's own vocabulary?

VI. To what extent does general excellence in composition depend upon possession of a large vocabulary?

VII. How does a student's ability in reading compare with ability of composition and knowledge of English words?

and such other factors as may seem relevant concerning age, sex, whether habitation is rural or urban, etc.

CHAPTER II

SOURCES OF DATA.

In the October 1934 issue of the A.T.A. Magazine, official organ of the Alberta Teachers' Alliance, there was published an appeal to the teachers of the province to obtain themes from their students in Grades VII to XII upon the topic "The Most Exciting Ride I Ever Had". The response to this appeal was most generous, and several hundred themes were received. From this source it was intended to form a selection of about five hundred subjects distributed as evenly as possible over Grades VII to XII inclusive and distributed between schools of as many different sizes as possible.

This ideal was not wholly obtained for various reasons. Few of the smaller school and practically none of the rural schools apparently teach Grade XII, and consequently most of the subjects in this grade were drawn from one source. However, since many of the students, after completing the earlier part of their secondary education had come to the city for their Grade XII, a fairly representative selection, it is felt, of the Grade XII population has been obtained.

The complete data that was desired from each child included the following:

I. Rating of composition according to Hudelson scale.

II. Analysis of fifty running words from the compositions for comparative word frequencies, using Thorndike's "Teacher's Word Book of 20,000 Words".

III. Otis Intelligence Quotient.

IV. Reading score according to the "New Stanford Reading Test".

V. Score on vocabulary test devised by the writer.

VI. Supplementary information such as:

(a) Age

(b) Sex

(c) Grade

(d) Language spoken at home

(e) Occupation of father

(f) Country of birth.

Among the several hundred essays submitted, there were comparatively few from Grade XI and none at all in this grade from city schools. To remedy this discrepancy, thirty eight students from an Edmonton school were asked to write the theme in the middle of February, about four months after the others had been written. This will perhaps bias slightly the results from this grade.

The high number of subjects from Grade VII is due to the fact that there was a misunderstanding as to what room subsequent tests should be given, with the result that they were given to the wrong class, necessitating an analysis, which would otherwise not have been done, of the compositions written by that class.

The original distribution of subjects by grades and schools is shown in Table I.

From this table we can see that of the whole population, the two cities of Edmonton and Calgary are together responsible for

TABLE I

DISTRIBUTION OF SUBJECTS.

School	Location	XII	XI	X	IX	VIII	VII	TOTAL
Bellevue	Bellevue	10	11	11	26	22		80
Camrose Practise	Camrose						33	33
Dalemead Rural H.S.	Dalemead		5	6	1			12
East Coulee	East Coulee	4	14	4				22
Garnea H.S.	Edmonton				32			32
Highlands	Edmonton					62	30	92
Hillock	Ranfurly		4	3	1	3	2	13
Lamont	Lamont		8	2	10	16	14	50
Mapova	Mapova					2	2	4
Merna	Sedgewick		8	4	2	4		8
Nacotown	Naco					6	1	7
Pincher City	Pincher Station			1		1	2	4
Redcliff	Redcliff	8	10					18
Strathcona	Edmonton	57						57
Sunniebend	Pibrock					5	2	7
Tofield	Tofield				11			11
Turin	Iron Springs	1		2	2		4	9
Victoria	Edmonton		38					38
Wallace	Shaughnessy		2	1	3	4	3	13
Western Canada	Calgary			30				30
TOTALS		80	100	64	88	125	93	550

two hundred and forty nine subjects, or 45.3% of the total. The reason for this is mainly that of expediency, to save the expense of time and money involved in sending small bundles of testing material to schools where there are only a few students, with the attendant uncertainty and irregularity of receiving the tests again for marking when wanted. However, it is partly due to the necessity of going to city schools for Grade XII pupils, the fact that few papers were received from Grade XI, and the necessity of adding an additional Grade VII class as mentioned above.

Two hundred and fourteen, or 38.9% of the essays analyzed were received from small town or village schools, including Bellevue, Camrose, East Coulee, Lamont, Redcliff, and Tofield. Eighty seven essays, or 15.8% of the essays came from rural schools, including Dalemead, Hillock, Mapova, Merna, Nacotown, Pincher City, Sunniebend, Turin, and Wallace.

Table I does not however represent the complete population studied. In the first place, many more compositions than those of the students included in this table were scored or rated. Also, students who were absent when the topic was assigned were present to take some or all of the subsequent tests, and similarly students present for the writing of the theme were absent in some cases for the other material. This means that the total population studied, even disregarding those who wrote the composition and nothing else, is somewhat larger than that indicated, whereas the population upon which complete data were obtained is somewhat smaller. This will be indicated where relevant.

CHAPTER III

SCORING THE THEMES

The themes were first of all scored by comparing them with the samples in the Hudelson English composition scale. This is a collection of compositions which are given values ranging from 2.00 to 9.5, the step or difference between each composition and the next in value being .5. The difference in ability from one composition to the next in the series is supposed to be equal all along the line, so that a rating of "6.0 represents exactly twice as much composition merit" as a rating of 3.0. The top two are the only compositions not written by school children.

However, it must be admitted that, though the investigator practised on the samples provided until his "systematic error" had reached a negligible amount, this rating is not wholly satisfactory. One hundred and forty of the essays shown were rated by a person other than the investigator. Moreover, despite the utmost care, the fact remains that this method of rating is not objective, and many factors, such as grammar, punctua tion, mastery of idiom, spelling, clarity, choice of language, maturity of style, and numerous others must be taken into account; and while one rater might place emphasis on one phase in forming his judgments, another person might lay stress on quite different factors. Also, the rating might be biased by such extraneous factors as neatness and writing, despite all efforts to the contrary.

Nor did the investigator find the scale wholly satisfactory. It seemed that the series did not contain sufficient samples and accordingly it was quite difficult to rate one that fell between two consecutive samples on the scale. Moveover, the samples are quite short in proportion to those obtained from Alberta students, and it was difficult to decide how far verbosity was a virtue or the reverse. Finally, the samples rated 4.5, 5.0,5.5, in the series seem rated too high in comparison with the others.

In compiling diagrams and calculating means, medians, norms, etc., the writer has taken into account all the themes he scored, regardless as to whether they are included in the populationas shown in Table I or not. This is done so that by having a larger population the results for Alberta may be more valid. However, in order that the results may be more homogeneous, only those themes which he himself scored have been considered. In making comparisons, deriving coefficients of correlation, etc., the rating of the Table I population alone has been considered, regardless of who did the scoring.

Five hundred and ten scores have been included in the analysis of the Hudelson ratings, the distribution according to grade and merit being indicated on Table II. In this Table are indicated the norms that the author of the test has derived, the means, medians, and Standard Deviations of our five hundred and ten cases, as well as the third and first quartiles (Q_3 and Q_1) and the Semi-Interquartile Ranges (s). Diagrams I to VI give the frequency polygons of the distributions from Grades VII to XII respectively,

TABLE II

DISTRIBUTION OF HUDELSON SCORES FROM GRADES VII TO XII, WITH RELEVANT DATA

Hudelson Score	XII	XI	X	IX	VIII	VII
8.1 - 8.2	2					
7.9 - 8.0			1			
7.7 - 7.8	2					
7.5 - 7.6	2					
7.3 - 7.4	5	1	1			
7.1 - 7.2	5	3		2		
6.9 - 7.0	1	1	2	1		
6.7 - 6.8	5	4	6	1	1	1
6.5 - 6.6	10	12	4	7	2	
6.3 - 6.4	6	17	28	11	8	3
6.1 - 6.2	2	14	38	12	17	4
5.9 - 6.0	1	16	30	16	20	7
5.7 - 5.8	2	6	12	11	24	7
5.5 - 5.6		1	15	11	9	3
5.3 - 5.4.	1	1	2	8	8	5
5.1 - 5.2					3	7
4.9 - 5.0			1	1	4	10
4.7 - 4.8			1	1	1	7
4.5 - 4.6					1	4
4.3 - 4.4					2	4
4.1 - 4.2						2
3.9 - 4.0						
3.7 - 3.8					1	1
3.5 - 3.6						1
TOTALS	44	76	141	82	101	66
Hudelson Norms	6.7	6.3	5.9	5.5	5.3	4.7
Means	6.791	6.437	6.089	5.970	5.809	5.238
Standard Deviations	0.612	0.376	0.400	0.454	0.450	0.68
Medians	6.700	6.300	6.150	6.013	5.879	5.214
Q_3	7.300	6.429	6.348	6.327	6.132	5.857
Q_1	6.467	6.037	5.928	5.691	5.668	4.829
s	0.417	0.196	0.210	0.318	0.232	0.514

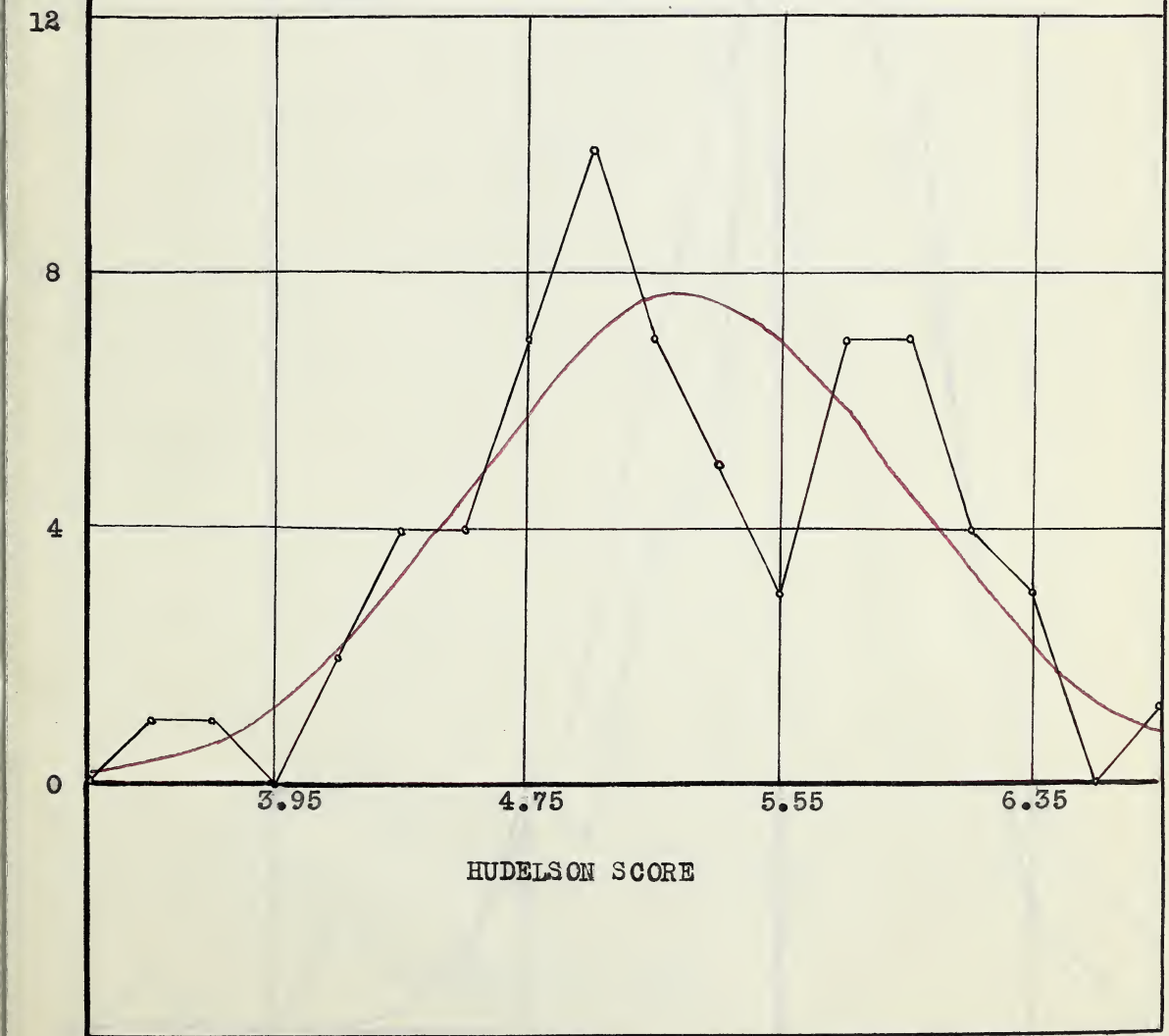
DIAGRAM I

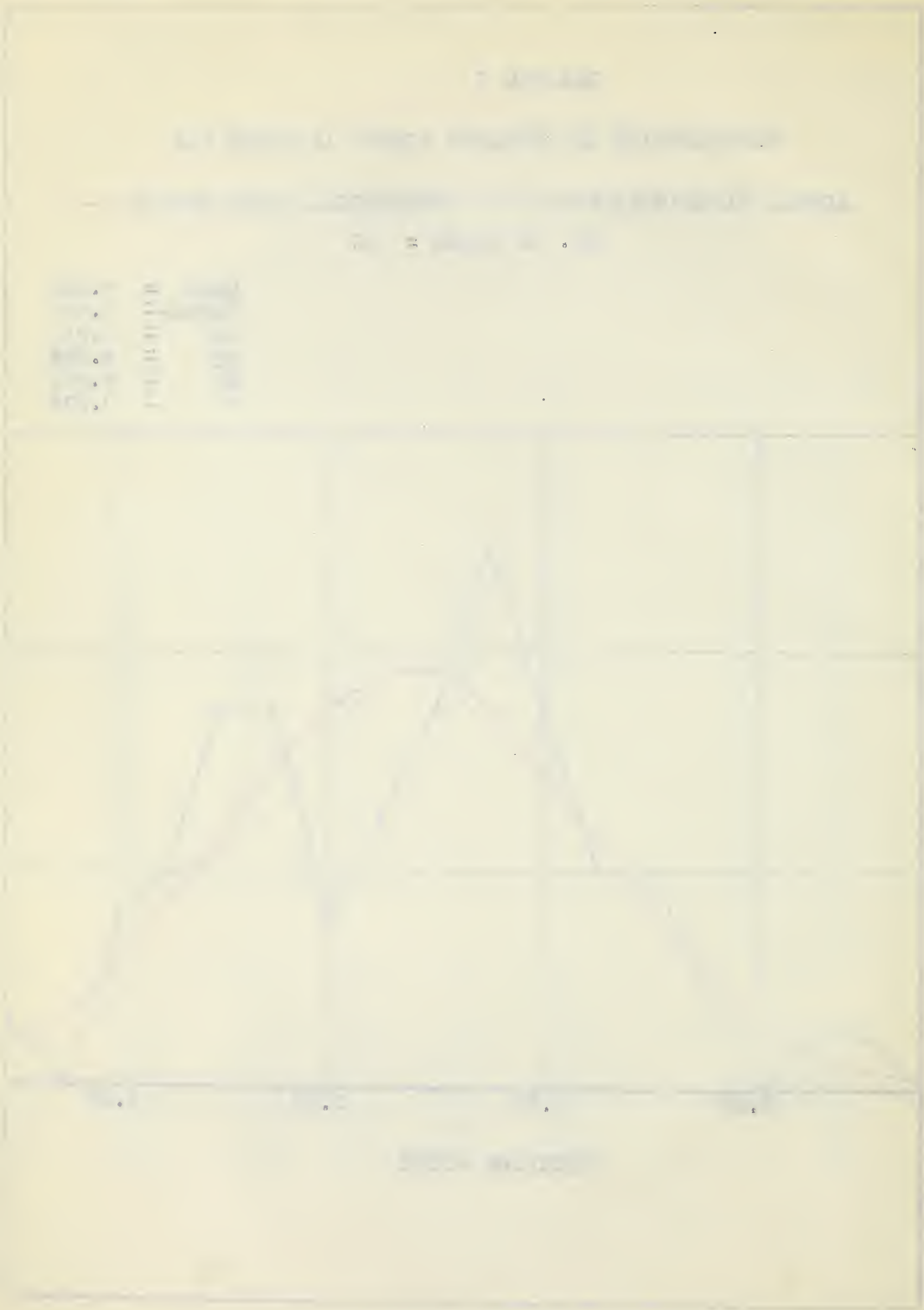
DISTRIBUTION OF HUDELSON SCORES IN GRADE VII

ACTUAL DISTRIBUTION——— THEORETICAL DISTRIBUTION———

NO. OF CASES = 66

Mean	=	5.238
Sigma	=	0.68
Q_3	=	5.857
Q_2	=	5.214
Q_1	=	4.829
s	=	0.514





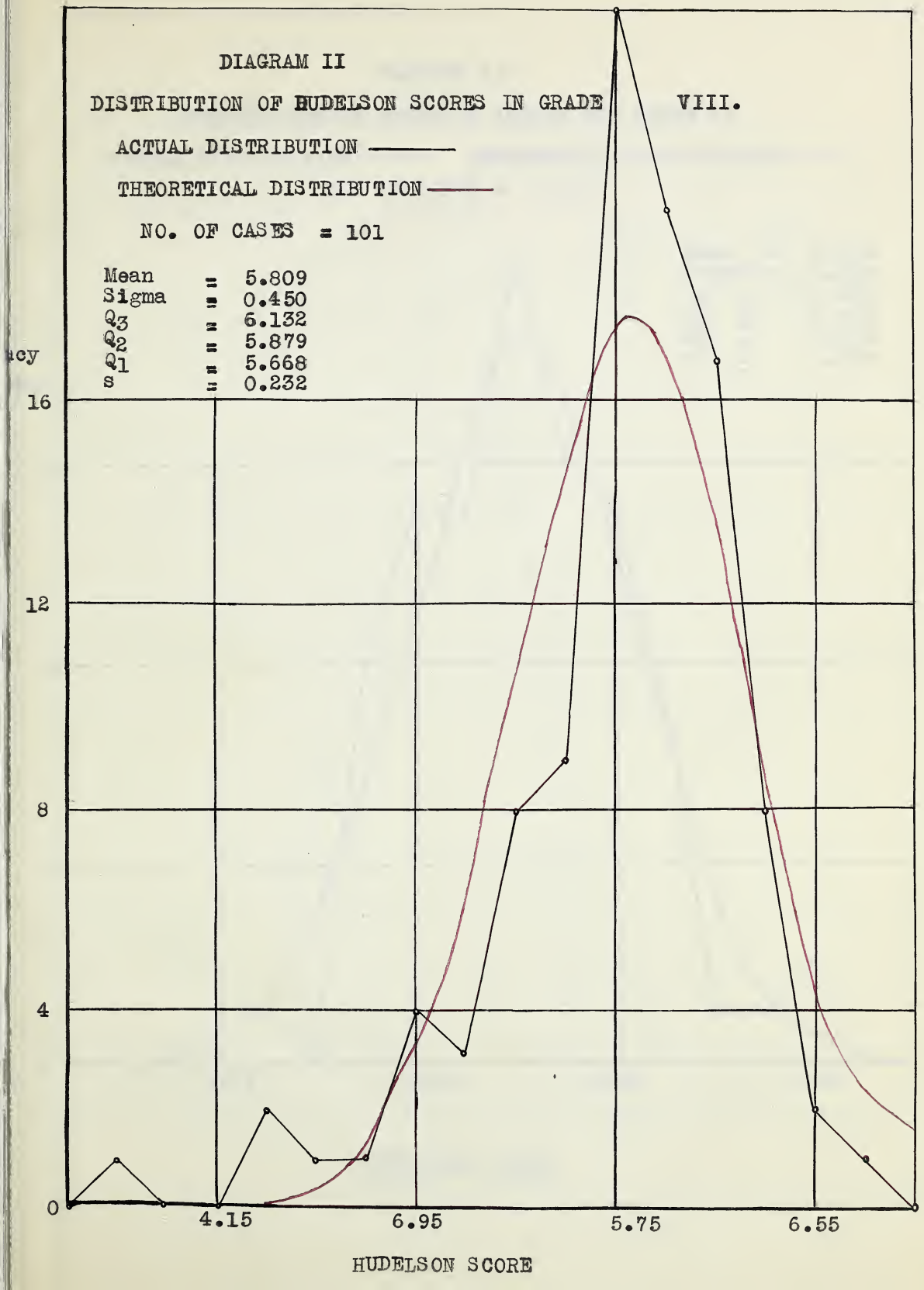


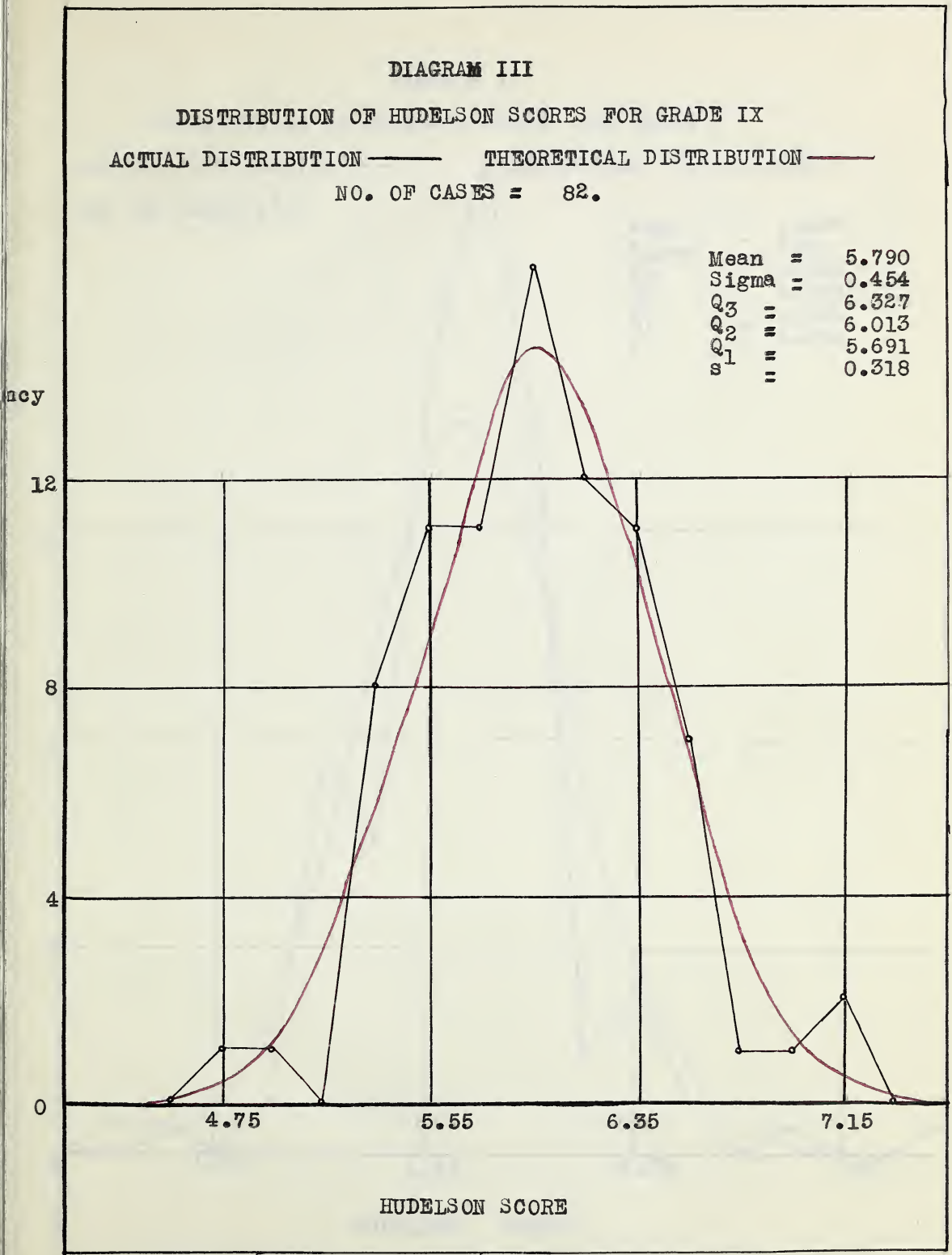
DIAGRAM II
DISTRIBUTION OF HUDELSON SCORES IN GRADE VIII.
ACTUAL DISTRIBUTION
THEORETICAL DISTRIBUTION
NO. OF CASES = 101
Mean = 5.809
Sigma = 0.450
Q3 = 6.132
Q2 = 5.879
Q1 = 5.668
s = 0.232
16
12
8
4
0
4.15
6.95
5.75
6.55
HUDELSON SCORE

DIAGRAM III

DISTRIBUTION OF HUDELSON SCORES FOR GRADE IX

ACTUAL DISTRIBUTION ——— THEORETICAL DISTRIBUTION ———

NO. OF CASES = 82.



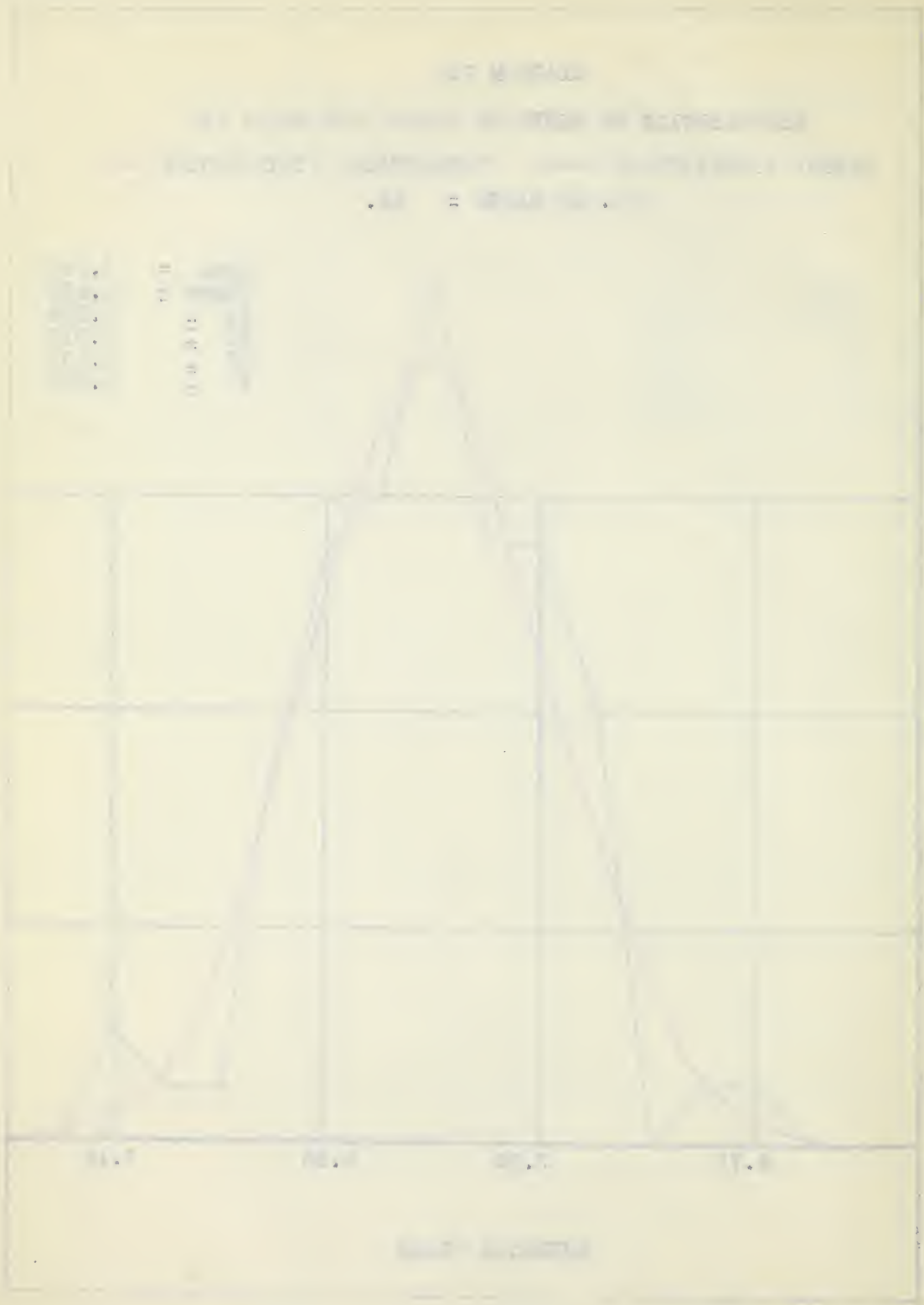
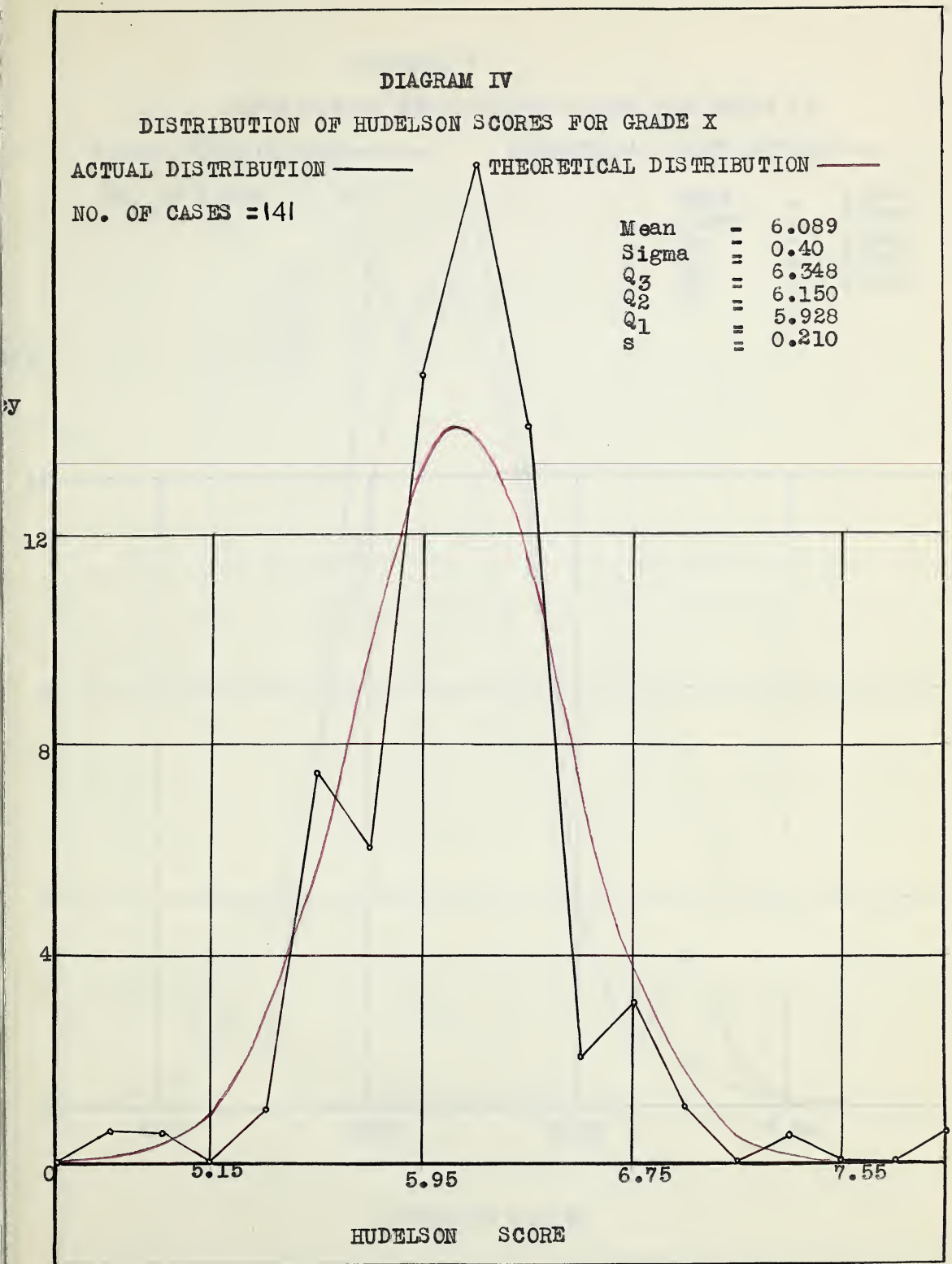


DIAGRAM IV

DISTRIBUTION OF HUDELSON SCORES FOR GRADE X



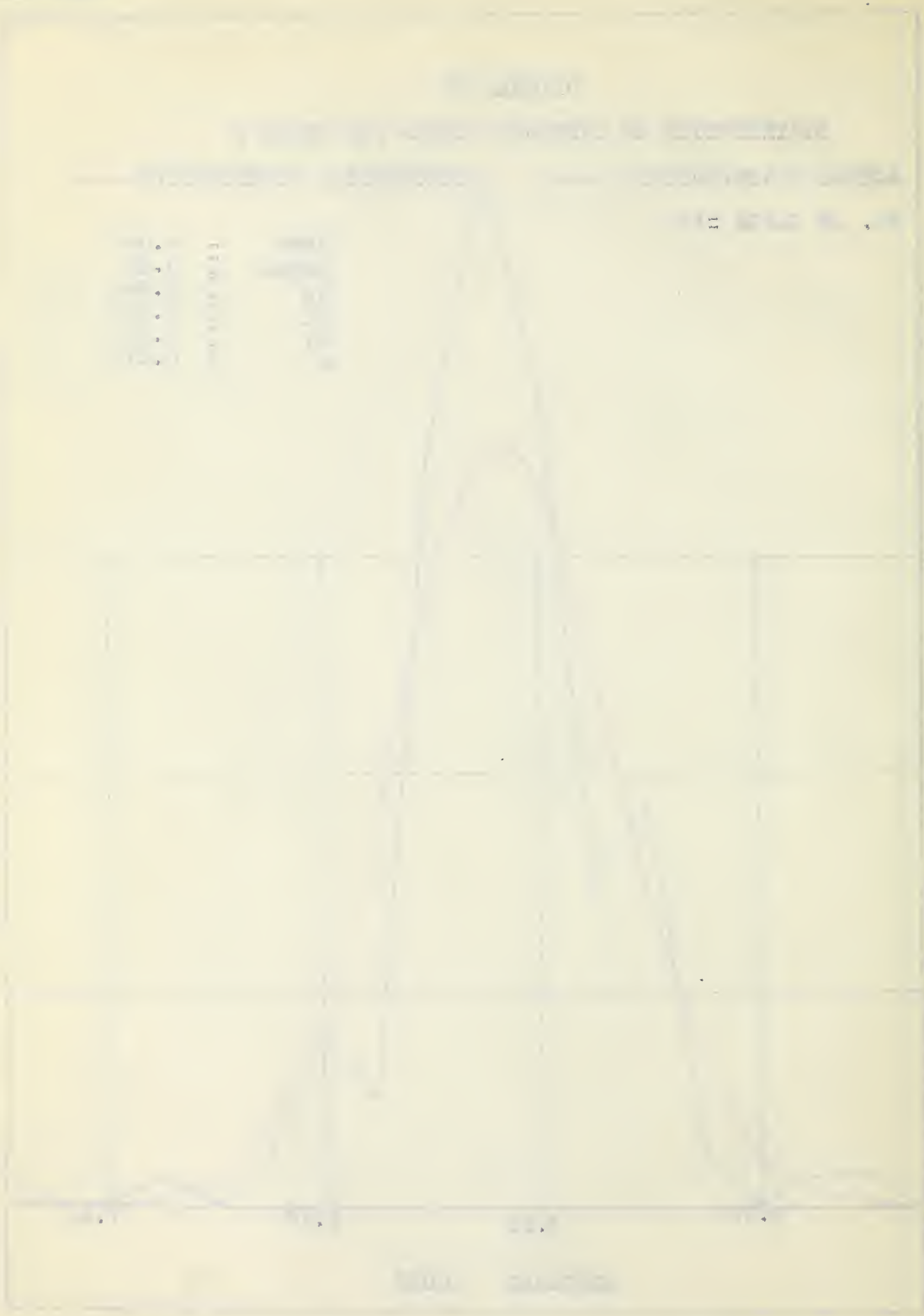


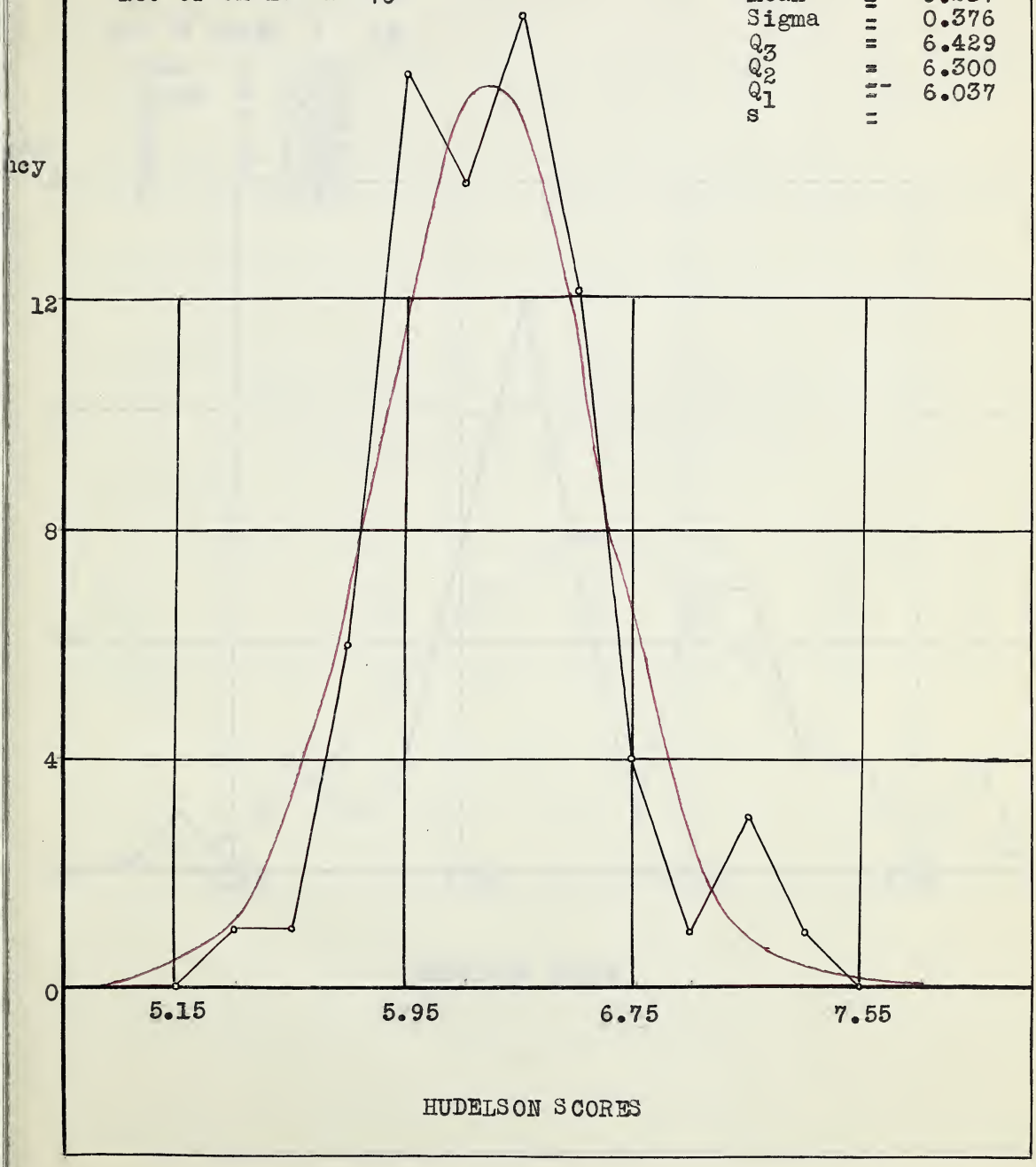
DIAGRAM V

DISTRIBUTION OF HUDELSON SCORES FOR GRADE XI

ACTUAL DISTRIBUTION ——— THEORETICAL DISTRIBUTION ———

NO. OF CASES = 76

Mean	=	6.237
Sigma	=	0.376
Q_3	=	6.429
Q_2	=	6.300
Q_1	=	6.037
s	=	



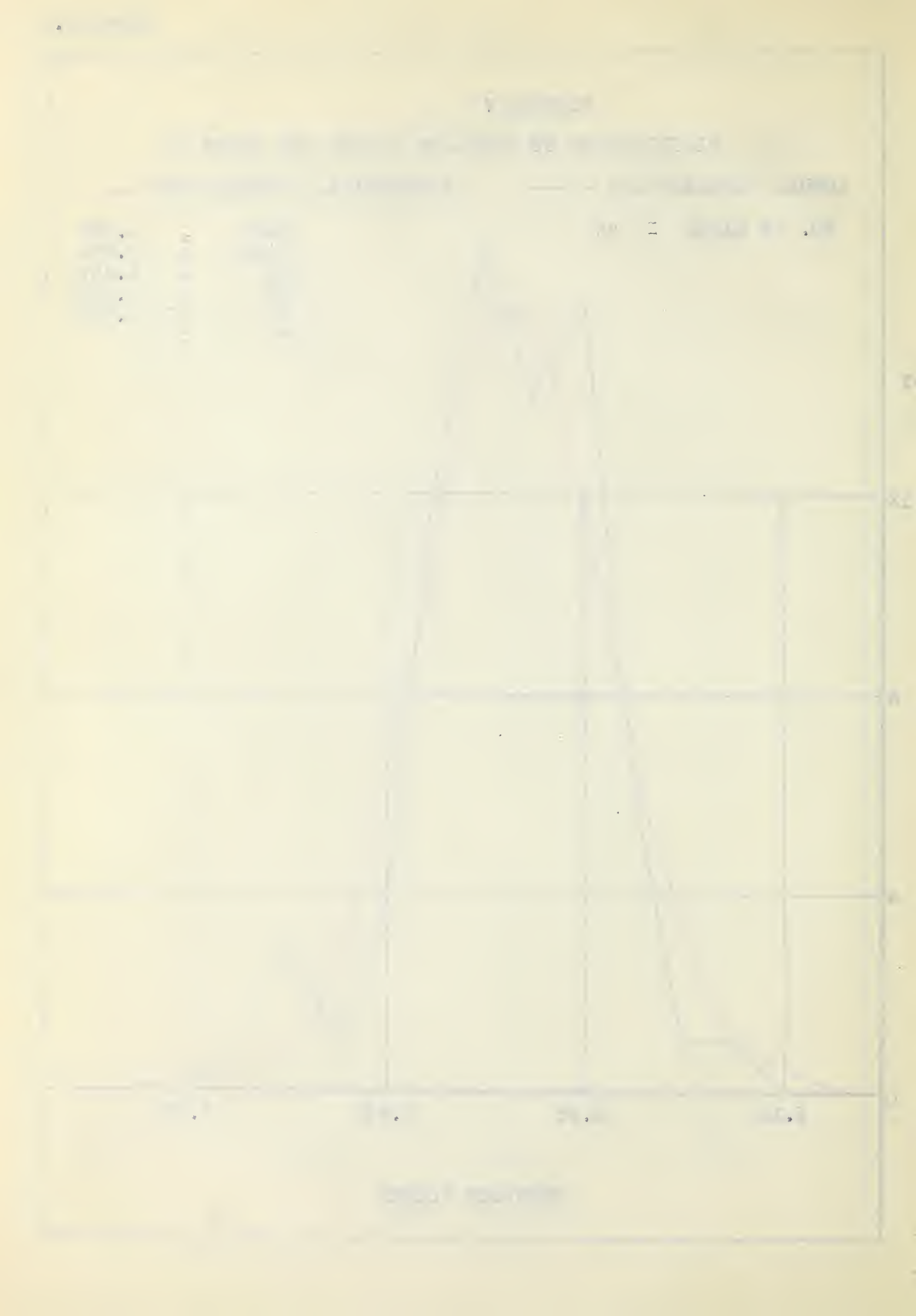


DIAGRAM VI

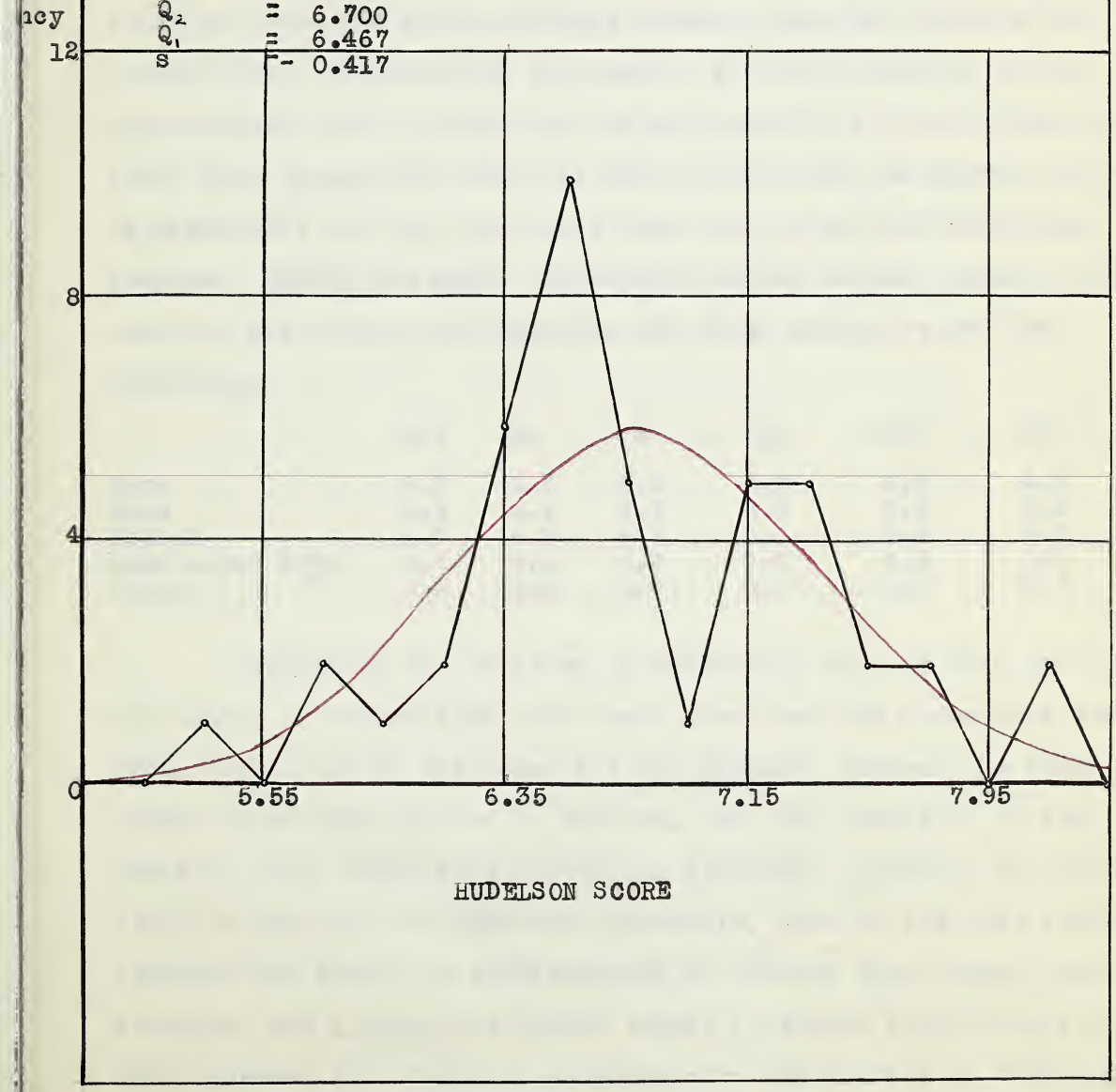
DISTRIBUTION OF HUDELSON SCORES IN GRADE XII

ACTUAL DISTRIBUTION ———

THEORETICAL DISTRIBUTION ———

NO. OF CASES = 44

Mean	=	6.791
Sigma	=	0.612
Q_3	=	7.300
Q_2	=	6.700
Q_1	=	6.467
s	=	0.417



with the data graduated by means of the normal curve according to the method illustrated on page 73 of Camp's "Elementary Statistics".

Hudelson does not tell us whether he derived his norms by means of the Medians or the Means, but this does not greatly matter, as in our data the correspondence between these two measures is fairly close. Remembering the sources of error inherent in the comparatively small population and subjectivity of the marking, it seems that though his norms are fairly applicable to Alberta students in Grades XII and XI, below that they are too low by increasing amounts. Taking the Means and Medians to one decimal place as Hudelson does his norms, and comparing the three scales we get the following:

	XII	XI	X	IX	VIII	VII
Norm	6.7	6.3	5.9	5.5	5.3	4.7
Mean	6.8	6.4	6.1	6.0	5.8	5.2
Median	6.7	6.3	6.1	6.0	5.8	5.2
Mean minus Norm	0.1	0.1	0.2	0.5	0.5	0.5
Median " "	0.0	0.0	0.2	0.5	0.5	0.5

Apparently the children in Grades VII to X in this province are better in composition than those from whom the norms were derived, especially as the Norms are for January, whereas the Alberta compositions were written in October, with the exception of half the Grade XI ones, which were written in February. Indirect corroboration to this will be furnished presently, when we see that Alberta students are better at comprehension in reading than certain American students, and <u>a priori</u> we should expect a certain relationship to exist between excellence of comprehension and ability in handling the vernacular in writing.

A comparison of the frequency polygons with the theoretical normal curves reveals a quite close correspondence in all cases except for Grades VII and XII. These grades had the lowest populations, the number of subjects being sixty-six and forty-four; and the discrepancies between observed and theoretical discrepancies is probably largely due to the paucity of data. However, it is interesting to observe that the sharp drop which occurs in both cases just above the middle of the range is strikingly similar to the drop that occurs just below the fifty percent mark in certain departmental high school examinations. Such discrepancies are well shown in a dissertation by Monroe MacLeod entitled "A Statistical Study of the Validity of Objective and Essay Examinations in History II". In that case the reason for the sudden drop seems to be that the examin ers are loathe to give a mark just below the passing mark, and shove it up or down a bit. Similar subjective factors no doubt enter here. In Grade VII the nadir of the depression occurs at the point 5.55 and in Grade XII at the point 6.95. In the opinion of the writer, the Hudelson value 5.5 is the point at which a composition changes from being definitely ppor to being tolerably good; and the value 7.0 marks the point at which a composition ceases to be immature, adolescent, or childish, and becomes representative of good mature writing. This is probably at least part of the explanation for these discrepancies.

In the case of Grade XI the scores are definitely skewed negatively from the ideal distribution. The quartile co-efficient

of Skewness, $(Q3 - 2Q2 + Q1)/s$, where the Q's are the quartiles and "s" the semi-interquartile range, is in this case —0.417. The Grade VIII scores are also somewhat skewed, but in the other direction. Here the measure of skewness is +0.232. In Grade X the scores clustertoo closely around the point just above the Hudelson score 6.0.

Considering the subjectivity of marking, the scoring seems to have been done reasonably well, especially in Grades VIII to XI inclusive. No doubt this is partly due to compensating errors, one theme being marked too low, and the next too high, and so on.

CHAPTER IV

ANALYZING THE WRITTEN VOCABULARY

In studying the importance of words actually used by students the compositions analyzed were those shown in Table I. From each theme fifty running words were chosen, starting about eight or ten lines from the beginning, and without reference to new sentences, paragraphs, etc. Then the comparative importance of each of these fifty words was found by reference to Thorndike's "Teacher's Word Book of 20,000 Words".

This book contains the twenty thousand words "found most frequently and widely in general reading for children and young people". Each of these words is classified according to its importance. The classifications run as follows: 1a, 1b, 2a, 2b-----5b, 6, 7,----20. Thus 1a means that a word is in the most important and most common thousand; and also that it is in the more common five hundred of that thousand. A word in the less common five hundred of the first thousand is classified as 1b and so on up to 5b. From there on the words are classified only to the nearest most common thousand; thus a rating of 14 implies that there are at least thirteen thousand, but not more than thirteen thousand nine hundred and ninety-nine words in the English language that are more common. In this study an additional classification, 20 + has been made for words that do not occur in the Word Book and are therefore less common than any of the twenty thousand words that do occur in it.

"Except for special reasons, separate entries are not make of plurals in s, plurals where y is replaced by ies, adverbs formed by adding ly, comparatives and superlatives formed by adding er, and est (including changes of y to ier and iest), or r, and st; verb forms in s, d, ed (including changes of y to ies and ied), and ing; past participles formed by adding n, and adjectives formed by adding n to proper nouns".

"Homographs like bear (the animal) and bear (carry) are counted together not because they really belong together but because by so doing the persons making the counts can make them much faster and with very much less strain of attention."

Thorndike warns us that this list cannot be considered perfect, despite the fact that over two hundred sources have been used, averaging over fifty thousand words each. "----the rarer words are not very accurately placed. Proper names have in general lower ratings than they should have.-----Some words recently introduced into books and magazines have lower credits than they should have. Words used in books about history and chemistry probably have somewhat higher ratings than they should have, and words used in books about music and athletics somewhat lower ratings than they should have. For certain reasons the credits for abbreviations like in., bbl., yd., etc., are probably too low and far less reliable than those for words.....Probably also there are some errors due to faulty marking, tabulating, or copying.....It is, however, enormously superior to any one person's judgement, and its occasional misleadingswill be outweighed probably fifty to one by its sub-

stantially correct guidance concerning the frequency and range of occurrence of words".

A study of the words used by Alberta students shows that often one word is used where another is intended, e.g. break for brake, new for knew, of for off, etc. In such cases the tabulation has been made from the word meant rather than actually written. Quite often also two words have been written where they should have been joined together. Examples of these are: motor cycle, street car, a while, up set. In each of such cases the composite words rather than the separate components have been tabulated. Occasionally also two words have been joined by the student, such as lightening-quick, fast-approaching, alright, etc. Where such combinations are found in the Word Book they are considered for this study as being one word; otherwise, they are considered as two and classified accordingly. Proper names, such as Jimmy, Idaho, Edmonton, etc., have been included where they occur in the Word Book, otherwise they have been skipped over, and considered as not having been written. This was thought preferable to classifying them as 20 +, as other words not in the Work Book have been classified, because in children's compositions a large number of place names and names of friends are mentioned, and it was felt that such procedure would bias the results too much in favor of the rarer words.

The comparative distribution of the words from classification 1b to 20 +, and from Grade VII to XII is shown in Table III. Since there are not equal numbers in each grade the frequencies have been

TABLE III

DISTRIBUTION OF VOCABULARY ACCORDING TO COMPARATIVE FREQUENCY OF WORDS. FREQUENCIES EXPRESSED AS PERCENTAGES.

	VII	VIII	IX	X	XI	XII
No. of subjects.	93	125	88	64	100	80
Word Importance.						
1a	82.69	81.00	78.10	79.00	77.22	76.12
1b	6.65	6.77	7.05	7.50	8.22	7.77
2a	2.54	4.00	3.27	4.09	4.46	4.57
2b	1.70	1.65	1.91	1.97	2.22	2.04
3a	1.31	1.04	1.57	1.72	1.22	1.45
3b	1.07	1.14	1.27	1.31	1.14	1.20
4a	.60	.69	.91	.65	.94	.92
4b	.47	.51	.86	.72	.72	.77
5a	.71	.72	.55	.27	.44	.60
5b	.15	.29	.39	.47	.38	.48
6	.54	.53	.73	.91	.72	.48
7	.19	.37	.36	.44	.46	.50
8	.13	.24	.32	.19	.02	.27
9	.22	.18	.09	.19	.16	.40
10	.19	.22	.23	.25	.28	.48
11	.11	.14	.27	.22	.18	.32
12	.17	.13	.14	.13	.12	.23
13	.06	.13	-	.16	.12	.07
14	.02	-	.07	.06	.16	.20
15	.06	.03	-	.03	.06	.10
16	.02	.02	.05	.09	.02	.07
17	.06	.03	-	.22	.10	.10
18	.04	.10	.09	.06	.06	.05
19	.02	.08	.07	.13	.04	.18
20	.09	.03	-	-	.08	.10
20 +	.31	.22	.36	.38	.38	.25

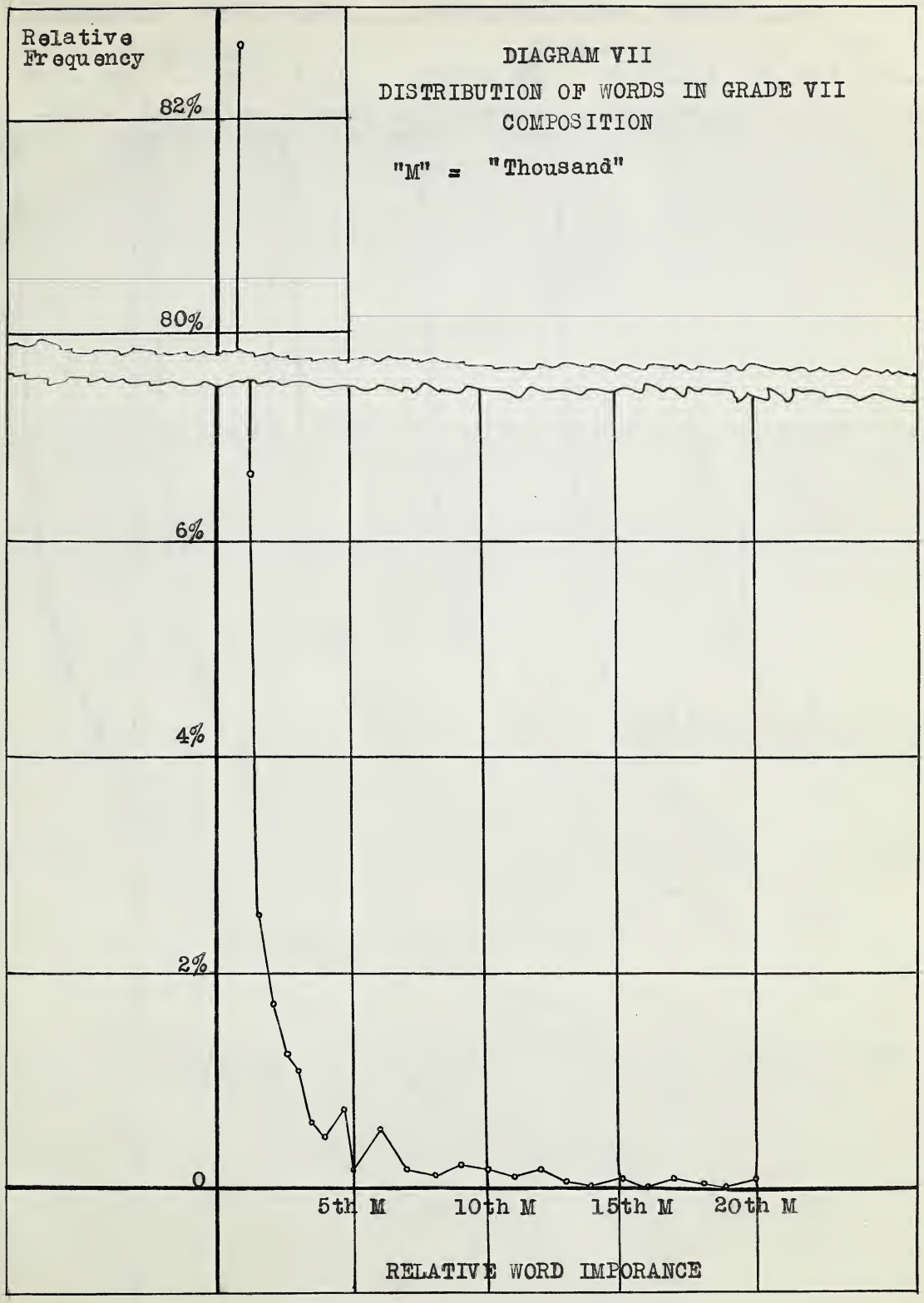
reduced to percentages to make them comparable with each other and to permit generalizations. From these results curves have been drawn in Diagrams VII to IX inclusive, using the classifications 1a to 20 as abscissae and the comparative frequencies as ordinates. For purposes of comparison the comparative frequencies in Grade VII from 1a to 20 have been plotted in Diagram VII on the usual Cartesian co-ordinates. This system does not reveal very much except that there is a very sudden drop at the beginning of the curve, and then the curve flattens out running more or less parallel to the horizontal axis. Accordingly, to get the whole curve on one sheet, and at the same time examine the lower part of the curve under a microscope, as it were, the curves have been plotted on Four Cycle Ratio Chart or Semi-Logarithmic graph paper. This has the advantage of revealing any trends by spacing out the points in the lower part of the curve, whereas such trends might easily be overlooked in a curve such as that in Diagram VII. However, in interpreting such curves, caution must be exercised, as the vertical distances between points on the lower half of such paper are greatly exaggerated. Naturally this is so when each deck or cycle represents ten times as great a vertical distance as the one immediately below it.

Diagram VIII shows the curves for Grades VII, VIII, and IX. The reason for plotting these curves on one sheet is partly because at the present moment such paper is practically unavailable. However, such a scheme is very serviceable for purposes of comparison. In order to avoid getting the curves tangled up with each other,

the vertical axis has been moved to the right two of the larger units for each successive grade. Diagram IX shows similar curves for Grades X, XI, and XII, with the tails of the preceding theoretical curves also indicated.

For each grade a smooth curve has been drawn by eye among the points. For various reasons this is preferable to the actual frequency polygon. In the first place there are so few words in the higher classifications that discrepancies easily occur, and the comparative frequency might easily be too high by one student using the same uncommon word two or three times in his fifty words. For instance, in Grade VII, with ninety-three students, all the words in the eighteenth thousand were contributed by one pupil, in the seventeenth and the thirteenth by two. Similarly, since words at such high altitudes, so to speak are quite rare, a classification might quite easily be skipped over, as has been done in Grade IX, with its eighty-eight students, in the thirteenth, fifteenth, seventeenth, and twentieth thousands. Incidentally, since no matter how far down one goes on semi-logarithmic paper one never reaches zero, points representing zero frequencies had to be omitted.

Another source of discrepancy lies in the topic itself. Since all subjects wrote on the most exciting ride they had ever experienced, we find the same words occurring over and over again. Thus the word 'pilot' occurs in the sixth thousand and that classification has consistently too great a frequency. 'Aeroplane (or airplane) is in the tenth thousand, and in every case but that for


Relative Frequency
82%
80%
6%
4%
2%
0
DIAGRAM VII
DISTRIBUTION OF WORDS IN GRADE VII COMPOSITION
"M" = "Thousand"
5th M
10th M
15th M
20th M
RELATIVE WORD IMPORANCE

Comparative Distribution of Words in Written Compositions

Actual Distribution - - - - - - Assumed Distribution ———

Grade VII —— Grade VIII —— Grade IX —— Grades X, XI, and XII ——

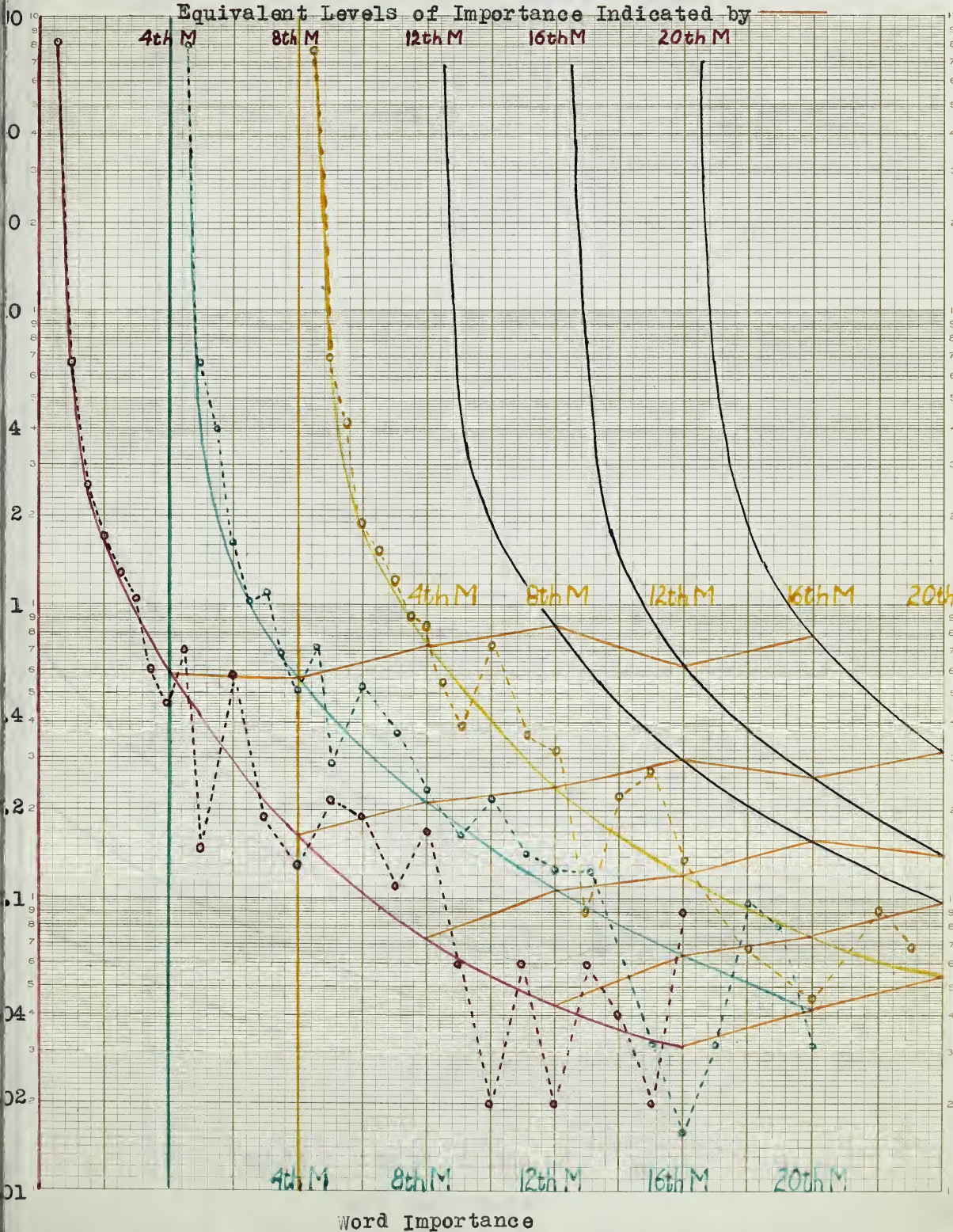


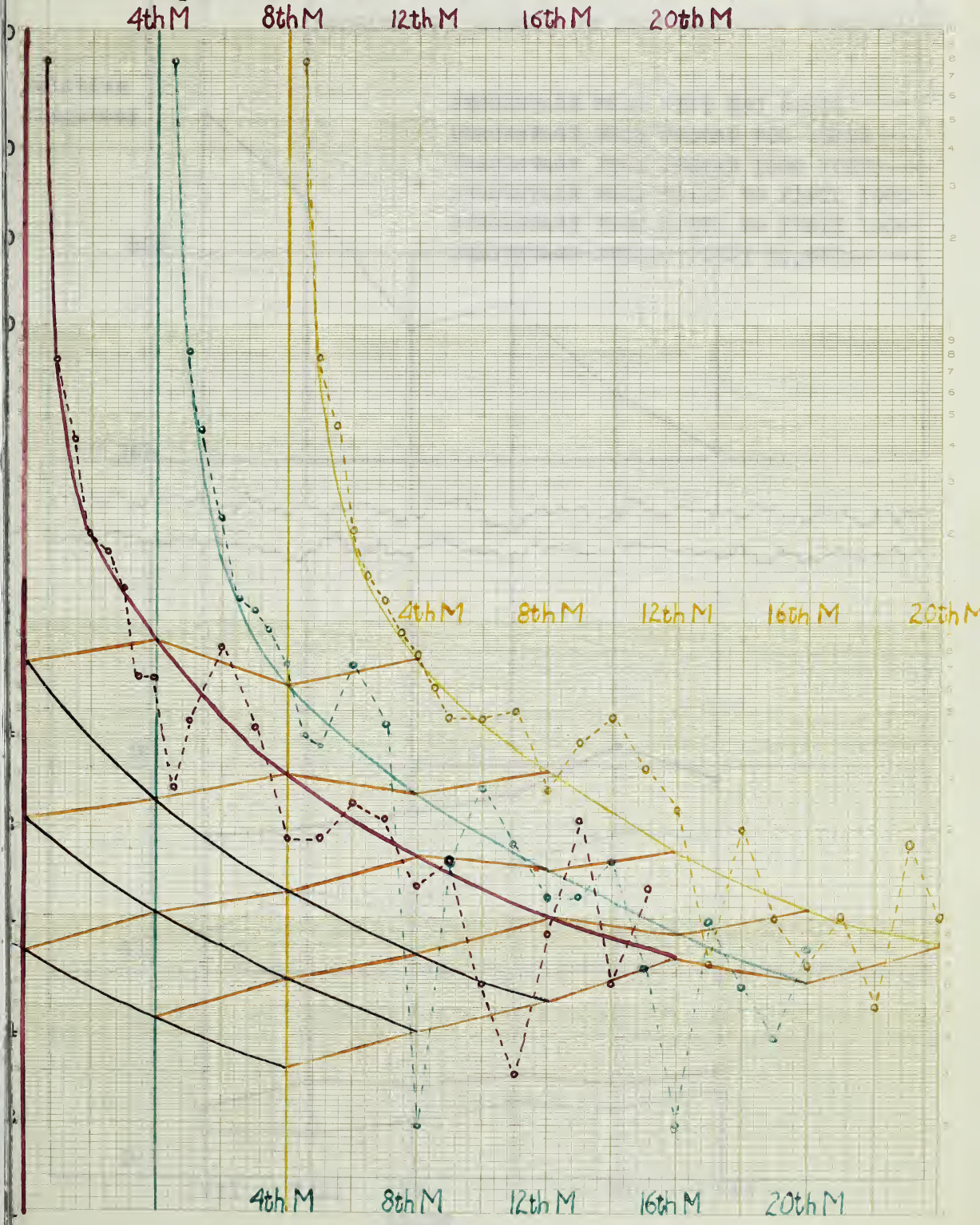
DIAGRAM IX

Comparative Distribution of Words in Written Compositions (continued)

Actual Distribution — Assumed distribution

Grades VII, VIII, and IX —— Grade X ——— Grade XI —— Grade XII —

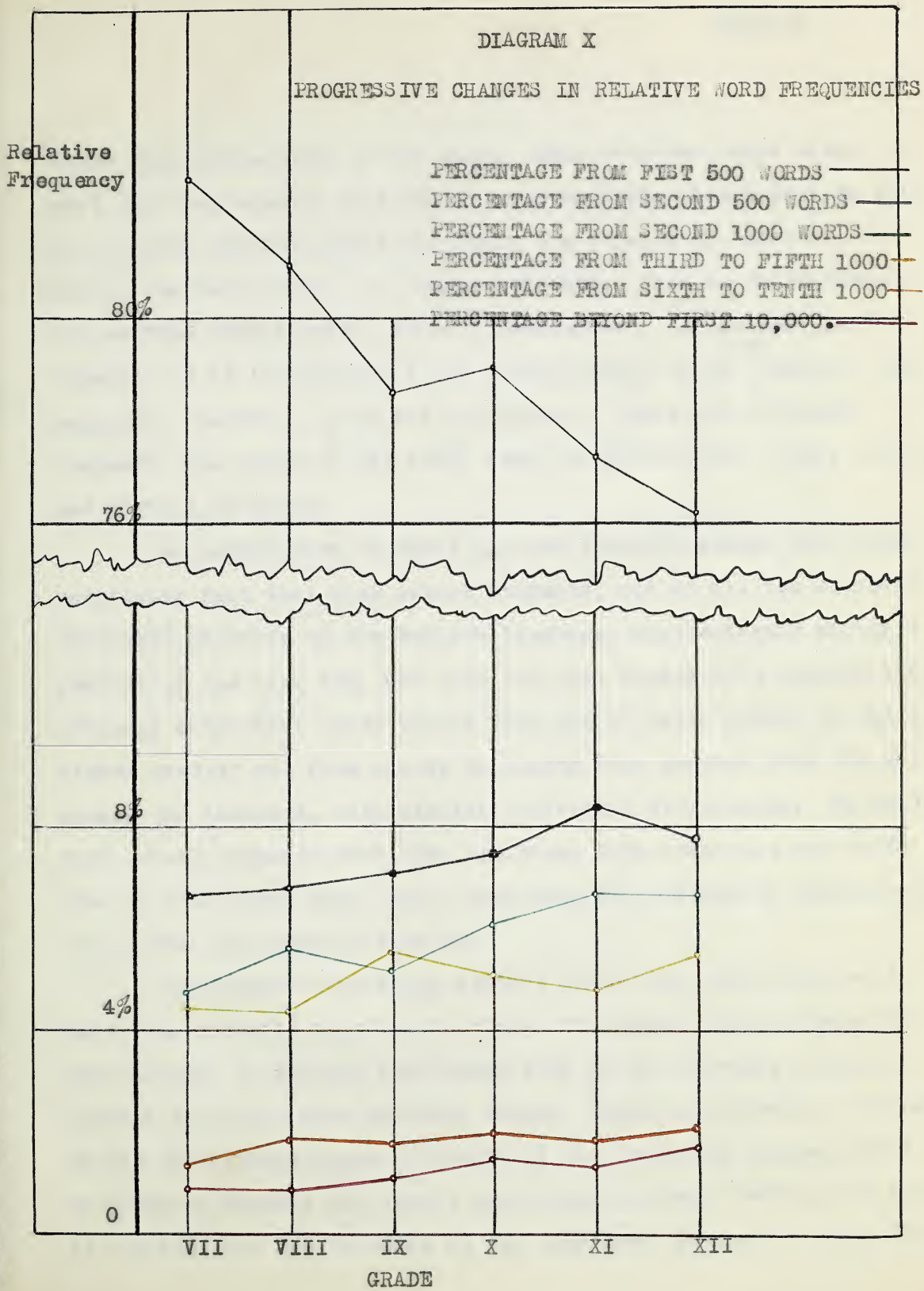
Equivalent Levels of Importance Indicated by —



Word Importance

DIAGRAM X

PROGRESSIVE CHANGES IN RELATIVE WORD FREQUENCIES



Grade XI the frequency is too high. Many children found their most exciting moments on a roller coaster (even though they do call it a rolley coaster) and accordingly the nineteenth thousand, in which 'coaster) occurs, is unduly affected. Thus in Classification 17, we find 'windshield'; in 18 'speedometer', 'propeller', and 'shaky'; in 19 'accelerator' and 'street-car'; in 20 'trailer' and goggles'. However, up to and including at least the fifteenth thousand, the topic of the essay seems to provide for a wide choice and variety of words.

An examination of Table III and derived diagram reveals the rather surprising fact that high school students, out of all the scores of thousands of words in the English language, choose around eighty percent of the ones they use from the most common five hundred, words, although individual cases varied from 94% in lower grades to 58% in higher grades; and from ninety to ninety four percent from the most common two thousand, with similar individual differences. No wonder high school students don't like spelling; they never use the words when they do learn them, and a knowledge of spelling is unnecessary in reading, or comparatively so!

The second surprising feature about this set of curves is that they indicate very little change from Grade VII to Grade XII. Acquisitions to written vocabulary seem to be extremely gradual, especially of the more uncommon words. There is an obvious increase in the theoretical number of words at the Twentieth thousand mark as shown in Diagram IX, and at corresponding lower levels, but when we realize that the increase at the twentieth thousand is only from

.03 to .08 percent, it becomes practically negligible.

In order to study any change in the comparative frequencies of words at different grade levels, Table III has been compressed by grouping certain frequencies together into the following:

Word Importance.	VII	VIII	IX	X	XI	XII
1a	82.67	81.00	78.10	79.00	77.22	76.12
1b	6.65	6.77	7.05	7.50	8.22	7.77
2	4.75	5.65	5.18	6.07	6.68	6.62
3 - 5	4.42	4.38	5.55	5.15	4.82	5.43
6 - 10	1.27	1.83	1.73	1.97	1.92	2.12
11 - 20 +	.91	.91	1.04	1.47	1.32	1.67

These results have been plotted in Diagram X, the abscissae representing grades and the ordinates representing comparative frequencies. This diagram shows that there is a definite and well-marked drop in the use of the first five hundred words, and a gradual increase in the use of words of lesser importance, which increase becomes less and less as higher and higher levels are reached.

These two factors,- meagreness of vocabulary at all educational levels, and extreme slowness of growth, have been noted by several previous investigators. G. C. Brandenburg conducted an investigation which is reported under the title of "Pyschological Aspects of Language" in the Journal of Educational Psychology, volume 9, pages 313 - 332, 1918. However, his test was more in the nature of a word recognition test. He chose two hundred words evenly spaced throughout the Webster Academic Dictionary, and each word in his test represented one hundred and forty in this dictionary. They were arranged in increasing order of difficulty, and the criterion of the child's knowledge of the word was his ability to

use it in a sentence. Two thousand pupils were given this test. They were drawn from sixty-eight classes in Grades II to XII inclusive in sixteen schools. These schools were located in towns scattered through Wisconsin, Iowa, Missouri, and Colorado, varying in population from twelve hundred to thirty thousand. The test was given in May. Part of his results were as follows:

GRADE	NO. OF CASES	AVERAGE VOCABULARY	GAIN OVER PREVIOUS YEAR
VII	300	11445	1832
VIII	255	12819	1374
IX	72	13504	685
X	71	15340	1736
XI	71	13974	
XII	41	14975	

If conditions and circumstances between 1918 and 1934, between Alberta and the American states mentioned above, are comparable, it is quite evident that there is little relationship between the size of a person's writing and his reading vocabulary.

We shall find this corroborated when we come to consider the Vocabulary Test the Alberta children answered.

F. P. OBrien conducted an investigation somewhat similar to the present one which he describes in "The Vocabulary of High School Students in Written Composition", Journal of Educational Research, volume 11, pages 344 - 350, 1925. He tabulated his results according to Thorndike's Teacher's Word Book. This goes only as far as the first ten thousand words in the "Teacher's Word Book of 20,000 Words", which is a revision of it. As far as it goes, it is substantially the same as the latter, although a few words have had their classification changed. He drew his subjects from fourteen

schools, which he divides into two Groups, Group I included five schools with enrollments ranging from one hundred and sixty to six hundred students, and Grades IX to XII inclusive. Group II took in nine schools with enrollments of fifty to one hundred and fifty pupils, and included Grades VII to XII, althought not every school had all these grades. Where we have used the arithmetical mean percentage, he used the median percentage. The median percentage of words chosen from the first one thousand is shown below.

Grade	Group I		Group II	
	No. of cases	Percentage of words in 1st 1,000	No. of cases	%age of words in 2nd 1,000
VII	184	91.2	89	90
VIII	214	91	86	90
IX	489	89	140	90
X	426	88	159	90
XI	343	87.7	114	89
XII	204	87.5	94	88
	1860		682	

OBrien sadly comes to the conclusion that "----the mere recognition of words gives no definite assurance that they will function in writing".

P. A. Witty and M. Fry, in an article entitled "Vocabulary Content of Compositions Written by College Students", Journal of Educational Research, volume 19, pages 135-138, 1929, discuss a similar investigation that they carried out on college students.

Five classes of upperclassmen and five classes of freshmen

were chosen from the School of Education and the College of Liberal Arts in the University of Kansas. There were altogether one hundred and seventy freshmen, and the same number of seniors and juniors. As in OBrien's experiment, the older edition of Thorndike's Word Book was used. In this edition not all the words were classified in even thousands, but that it not important. His results were as follows:

Word Classification	Freshmen	Juniors	Seniors
1	85.35	84.50	84.30
2	5.47	6.28	5.81
3	3.44	3.59	4.06
4	1.70	1.78	2.12
4.001 - 5.144	1.77	1.42	1.78
4.145 - 10	-	-	-
10 +	2.51	1.82	1.70

"The teacher of English must not expect training in word recognition to transfer to a large extent to word usage in composition writing". That makes it practically unanimous.

P. M. Symonds and Baldwin Lee have also studied the subject in somewhat the same manner, using the earlier edition of the Word Book, but as their results are not readily comparable with such as given above, they are omitted. Their findings are outlined in "Studies in the Learning of English Expression", Teacher's College Record, volume 31, pages 50 - 58, October 1929.

Table IV is a synthesis of the results obtained in this investigation (classifed under A); those which OBrien found (entered under BI and BII for Groups I and II respectively); and those obtained from college students by Witty and Fry (classified under C).

TABLE IV

WORD DISTRIBUTION ACCORDING TO GRADE AND FREQUENCY

FREQUENCIES EXPRESSED IN PERCENTAGES

Grade		1	2	3	4	5	6 - 10	10-
VII	A	89.34	4.75	2.38	1.07	.86	1.27	.93
	BI	91.2						
	BII	90						
VIII	A	87.77	5.65	2.18	1.10	.99	1.54	.91
	BI	91						
	BII	90						
IX	A	85.15	5.18	2.84	1.77	.94	1.73	1.05
	BI	89						
	BII	90						
X	A	86.50	6.07	3.03	1.37	.74	1.98	1.47
	BI	88						
	BII	90						
XI	A	85.44	6.68	2.36	1.66	.82	1.58	1.32
	BI	87.7						
	BII	89						
XII	A	83.87	6.62	2.65	1.69	1.08	2.13	1.68
	BI	87.5						
	BII	88						
Freshmen	C	85.35	5.47	3.44	1.70	1.55	.22	2.51
Juniors	C	84.50	6.28	3.59	1.78	1.24	.18	1.82
Seniors	C	84.30	5.81	4.06	2.12	1.56	.22	.170

Since the fifth classification did not exactly include the fifth most common thousand, these were found by a method of straight line interpolation, and the difference was classified under the head 6 - 10. This is probably not strictly accurate, but close enough for practical purposes. The Table shows us that regardless of time, place, or educational attainments, students choose from eighty-four to ninety percent of their words from the most common thousand, four to seven percent from the second most common thousand, two to four percent from the third thousand and so on. Alberta students seem to choose somewhat fewer from the first thousand than do other students of similar educational standing, which of course means that they use more words that are not so common than do others.

CHAPTER V

ANALYZING THE RECOGNIZED VOCABULARY

As mentioned in Chapter II, besides a study of the written vocabularies of the students, a vocabulary test was also given. This was derived from the Word Book and modeled after the Thorndike Test of Word Knowledge. The latter was designed for Grades IV to X, and as our subjects were in Grades VII to XII, it was desired to have one somewhat more difficult, which was accomplished by choosing more of the uncommon words than Thorndike has.

The test was drawn up on the multiple-choice form, so that for each test word the student had to find the best synonym from among five possible answers. The words were chosen and arranged according to decreasing frequency and importance in the language as measured by Thorndike's Word Book. The words were chosen according to the following scheme.

Five	words	from	each	of	the	first	four thousand
Six	"	"	"	"	"	next	nine "
Five	"	"	"	"	"	"	four "
Two	"	"	"	"	"	last	three "

At each level words were chosen so as to avoid inflections of more common words, such as 'knew', 'brought', etc., with one or two exceptions that will be pointed out below. Words beginning with prefixes, such as 'unseen', 'transcontinental', etc., were also avoided. The words were chosen from all through the book.

The correct response word was judged by three criteria. In

the first place, it had to be one for which a synonym could be found in a standard dictionary. For this purpose Annandale's "Concise English Dictionary" was used. In the second place the synonym had to be at least as common as the stimulus word itself, preferably more so. Finally, the stimulus word had to be capable of definition in one or at most two words. Where any one of these criteria could not be satisfied, the word was discarded and another chosen.

The position of the word was chosen at random. There were five possible positions. Starting at the beginning of the test, a column of figures from Wentworth and Hill's Logarithmic Tables was chosen and the last digit of the logarithm noted. When a 1 or 2 occurred, the response word was placed in the first position; where a 3 or 4 occurred, in the second position and so on.

Of the possible response words, particular attention was paid to one other besides the correct one. A word analogous in form or meaning; or suggestive in some way or other was used where possible. Thus for words similar in form we have 'precious' as an alternative for 'precarious', 'parrot' for 'parapet', 'lapel' for 'ladle', etc. For analogies in meaning, there is 'companion' for 'accompany', 'hatred' for 'rage', 'significant' for 'signify'. These occur mainly in the first part of the test. In the last part, this indirect suggestion was more subtle. Thus for No. 60, 'Boulder' is somewhat similar to 'bowler' and as the latter is a hat, the word was included in the possible choices. Similarly in No. 46, 'averse' suggests 'reverse', so 'opposite' was used. In No. 99 'jejune' resembles 'jeune' so the word 'young' was included.

The position of the analogous response word was also chosen at random. After the correct response had been filled in, there still remained four possible positions for the analogous word. Each suit in a deck of card was arbitrarily given a position and by shuffling and cutting the deck, the analogous words were located.

The remaining possible response words were chosen haphazardly being of about the same degree of difficulty as the stimulus word. Thorndike in his test has made all the possible responses to be the same part of speech, such as noun, adverb, verb, etc., as is the stimulus word. In this test, as long as the word looked as if it might be the same part of speech, it was considered satisfactory.

The following is the final form of the test, with the positions of the correct response word indicated in each case. Thirty minutes were allowed for its completion.

VOCABULARY TEST

Write your name plainly.....Grade.....
School.....Post Office.....
Birthday .....Age Last Birthday.....Sex.....
Country of Birth.....Language Spoken at Home.....
Occupation of Father.....

- - - - - - - - - -

Look at the first word in line A. Find the other word in the line that means the same, or almost the same. Write its number on the line at the edge of the page. Do the same with all the other lines. The first few lines show you how to do it. Do as many as you can.

A rapid 1. warm 2. moist 3. quick 4. near 5. now 3
B head 1. doorway 2. part of body 3. speed 4. hat 5. book 2
C cattle 1. barn 2. farmer 3. stove 4. cows 5. car 4
D demand 1. desire 2. ask for 3. shut 4. turn 5. open 2

- - - - - - - - - -

BEGIN:

1. ask 1. answer 2. tell 3. question 4. repair 5. reply 3

2.	gather	1. got 2. distribute 3. collect 4. goatherd 5 earn	3.
3.	Notice	1. inform 2.see 3.use 4. accept 5.sight	3.
4.	travel	1.carry 2,journey 3.refresh 4.train 5.approve	2.
5.	quiet	1.still 2.warm 3.dark 4.inquire 5.moist	1.
6.	final	1.find 2.autumnal 3.joyful 4.weary 5.last	5.
7.	immediate	1.remote 2.comfortable 3.important 4.close 5.necessary	4.
8.	beneath	1.neither 2.around 3.under 4.reddish 5. open	3.
9.	courage	1.courtesy 2.bravery 3.wisdom 4.discretion 5.love	2.
10.	accompany	1.attend 2.companion 3.remove 4.place 5. desire	1.
11.	gradual	1.stiff 2.winding 3.green 4.sudden 5. by degrees	5.
12.	pony	1.flower 2.island 3.small horse 4.large dog 5.cart	3.
13.	nearly	1.surely 2.almost 3.next 4.never 5.at last	2.
14.	rage	1.anger 2.hatred 3.friendship 4.excitement 5.revenge	1.
15.	glance	1.replace 2.decide 3.eyes 4.look quickly 5.open	4.
16.	hinder	1.remain 2.assist 3.prevent 4.below 5.eat	3.
17.	brilliant	1.shining 2.useless 3.fluent 4.green 5.bubbling	1.
18.	formerly	1.perhaps 2.soon 3.form 4.before 5. orderly	4.
19.	apparent	1.required 2.half open 3.clear 4.powerful 5. father	3.
20.	inhabit	1.occupy 2.habit 3.cultivate 4.home 5.forecast	1.
21.	signify	1.firefly 2.significant 3.write 4.mean 5.forecast	4.
22.	prairie	1.a grassy plain 2.garden 3.farm 4.desert 5.prepare	1.
23.	certainly	1.amplify 2.probably 3.hopefully 4.surely 5.naturally	4.
24.	chaff	1.wheat 2. husks 3.throw 4.board 5.staff	2.
25.	equip	1.very fast 2.strike 3.paint 4 provide 5.plant	4.
26.	humility	1.kindness 2.humbleness 3 envy 4.mantle 5.humidity	2.
27.	pilot	1.attorney 2.soldier 3.pile 4.guide 5.polite	4.
28.	uneven	1.rough 2 hard 3.hopeless 4.undone 5.accidental	1.
29.	successor	1.forefather 2.remainder 3.king 4.follower 5.guardian	4.
30.	devout	1.religious 2.handsome 3.morbid 4.resilient 5.hungry	1.
31.	exempt	1.empty 2.careless 3.free 4.peaceful 5.pale blue	3.
32.	hale	1.clever 2.lovable 3.superfluous 4.icy 5.healthy	5.
33.	hoist	1.turn 2.raise 3.hurt 4.determine 5.clasp	2.
34.	swerve	1.speed up 2.upset 3.stop 4.turn aside 5.sweat	4.
35.	panic	1.terror 2.pane 3.auditorium 4.heat 5.open.	1.
36.	award	1.hope for 2.away 3assign 4.request 5.shoot	3.
37.	uncouth	1.cruel 2.silly 3.uncover 4.bold 5.awkward	5.
38.	sunder	1.separate 2.loud 3.refulgent 4.under 5.sudden	1.
39.	precarious	1.precious 2.prominent 3.uncertain 4.probable 5.execrable	3
40.	dexterity	1.majesty 2.dextrose 3.tallness 4.monkey 5.expertness	5.
41.	brag	1.boast 2.orate 3.fool 4.lizard 5.rag bag	1.
42.	brandish	1.shut up 2.radish 3.flourish 4.hide 5.grey	3.
43.	feud	1.decision 2.advice 3.feudal 4.quarrel 5 friend	4.
44.	incorporate	1.business 2.improve 3.example 4.corporal 5.unite	5.
45.	kit	1.kitchen 2.outfit 3.nadir 4.skunk 5.equitable	2.
46.	averse	1.anxious 2.unwilling 3.reversal 4.poem 5.opposite	2.
47.	jumble	1.juniper 2.confusion 3.celerity 4.team 5.aloofness	2.
48.	formulate	1.state precisely 2.relation 3.boil 4.pair 5.conform	1.
49.	parapet	1.parrot 2.branch 3.order 4.dead 5.rampart	5.
50.	prowl	1.steal 2.wolf 3.shoot 4.inhibit 5.rove	5.
51.	eject	1.conjecture 2.exhibit 3.interrupt 4.expel 5.surrender	4.

52. holster 1.brick 2.halt 3.pistol-case 4.aborigine 5.confront 3.
53. sheen 1.shall not 2.brightness 3.foliage 4.purple 5.yarn 2.
54. ladle 1.lift out 2.lapel 3.stove 4.coat-hanger 5.lucid 1
55. irresolute 1.resolve 2.evanescent 3.tardy 4.climbing 5.undecided 5.
56. gusty 1.machine gun 2.soporific 3.mournful 4.tempestuous 4.
57. palfrey 1.balcony 2.horse 3.pallid 4.warrior 5.romance 5.digusted 2.
58. spherical 1.round 2.conical 3.solid 4.extreme 5.outstanding 1.
59. avalanche 1.efflorescence 2.windfall 3.glacier 4.warship 5.landslide 5.
60. bowlder 1.braver 2.excess 3.stone 4.hat 5.reformed 3.
61. chum 1.cosy 2.pal 3.chapter 4.minnow 5.child 2.
62. halve 1.irrigate 2.resilient 3.own 4.despair 5.bisect 5.
63. gentility 1.generosity 2.gentile 3.refinement 4.flattery 5.posterity 3.
64. avaricious 1. covetous 2.voracious 3.impressive 4.mighty 5.crumbling 1.
65. enigma 1.steam engine 2.decision 3.inspiration 4.clarity 5.puzzle 5.
66. fiscal 1.fishy 2.monetary 3.interesting 4.dormant 5.dead 2.
67. garner 1.succed 2.guarantee 3.place 4.replenish 5.gather up 5.
68. immutable 1.mysterious 2.merciful 3.invariable 4.mute 5.diseased 3.
69. immure 1.reside 2.verbose 3.laugh 4.manicure 5.confine 5.
70. turnpike 1.roadside 2.flower 3.toll-gate 4.piker 5. pathos 3.
71. cynicism 1.disintegration 2.wreckage 3.philosophy 4.transpiration 5.contempt 5.
72. auspices 1.hospital 2.prophecies 3.protection 4.prhhibition 5.liquids 3.
73. coyotes 1.gophers 2.prairie wolves 3.anteaters 4.hunters 5.environs 2.
74. infuriate 1.distill 2.infiltrate 3.synthesize 4.express 5.enrage 5.
75. vibrant 1.vibrous 2.tremulous 3.experienced 4.acme 5.interminable 2.
76. dyke 1.destroy 2.desperation 3.ditch 4.boy 5.marsh 3.
77. avowal 1.plenty 2.aspect 3.devour 4.mortify 5.declaration 5.
78. morose 1.ecstatic 2.sullen 3.tomorrow 4.cryptic 5.forgotten 2.
79. halibut 1.pigeons 2.fern 3.flat fish 4.halcyon 5.anterior 3.
80. momentum 1.extrovert 2.catalyst 3.momentary 4.biological 5.impetus 5.
81. avidity 1.eagerness 2.vision 3.zenith 4.surplus 5.aperture 1.
82. invulnerable 1.torrid 2.merciless 3.unassailable 4.inviolate 5.remarkable 3.
83. clandestine 1. candlestick 2.vernal 3.indubitable 4.secret 5. profane 4.
84. prescience 1.osprey 2.pathology 3.conscience 4.pressure 5. foreknowledge 5.
85. corpulent 1. fat 2.indigenous 3.corpse-like 4.psychic 5.reversible 1.
86. manacle 1.shackle 2.orthodoxy 3.barbarism 4.man-killer 5.hormone 1.
87. scutcheon 1.inebriation 2.spearhead 3.nonésense 4.shield 5.scuttle 4.
88. mania 1.pathology 2.partiality 3.liquid 4.remainder 5.madness 5.
89. scurrilous 1.cumulous 2.itchy 3.dynamic 4.abusive 5.considerate 4.
90. fulminate 1.fool proof 2.fulfil 3 explode 4.external 5.auroral 3.
91. remedial 1.personal 2.curative 3.medium 4.myopic 5.polite 2.
92. silvan 1.forest-like 2.silvery 3.permenent 4.comparative 5.endocrine 1.
93. oscillation 1.vibration 2.kiss 3.quarrel 4.distillation 5.amplitude 1.
94. sapient 1.economic 2.atomic 3.reprehensible 4.sage 5.sirupy 4.
95. refraction 1.unified 2.bending 3.fabrication 4.hysterics 5.pundit 2.
96. recuperate 1.recover 2.culpable 3.respiratory 4.customary 5.murder 1.
97. quandry 1.homicide 2.gentry 3.quarry 4.turn-out 5.predicament 5.
98. ostler 1.surgeon 2.tappet 3.stable-boy 4.hustler 5.façade 3.
99. jejune 1.young 2.uninteresting 3.empirical 4.octopus 5.clarifying 2.
100. halbert 1.kinsman 2.avuncular 3.halibut 4.prehensile 5.battle-axe 5.

Before the test was given to all the students, it was first tried out on three Edmonton classes, including one in Grade VII, one in Grade IX, and one in Grade XII. These classes included, thirty-four, twenty-nine, and twenty-four students respectively. The test seemed to give satisfactory distributions and be suitably graded for difficulty, the respective medians being 42.9, 67.5, and 78.3. Two slight changes only were made. In the original draft, No. 13, 'nearly' had been erroneously printed as 'nearby.' However, 'next' was considered as the correct response, and the tests marked accordingly. In No. 30, the original fifty response for 'devout' was 'devoted' which apparently was too close to the meaning of the stimulus word, as many chose it in preference to the correct response 'religious'. Accordingly, it was changed to 'hungry'.

In order to test the test, a number of students in Grades VII, VIII, and IX also answered Form A of the Thorndike Test for Word Knowledge, which as mentioned above, is very similar in form. In all, one hundred and eight students wrote the two tests, but for comparative purposes one hundred and seven were considered. On the above test one student had obviously guessed at every word, as his answers were all 2's and 3's and his total score only 19, whereas by the operation of chance alone he should get a score of 20. The usual product - moment method of deriving a co-efficient of correlation was utilized, with the following results:

	Thorndike's Test	Our Test
Mean	53.655	48.330
Standard Deviation.	11.78	11.60

$r = +0.91$

$P.E._r = 0.0106$

The co-efficient of correlation between this test is very high, with a very low Probable Error. Accordingly we may consider it to have a high degree of reliability. It is noteworthy that the Standard Deviations in these two tests for the population tested is almost the same. To compare the means, we must add 3 to the Thorndike' Mean, as that much is subtracted from each score to get the final score, probably to make the norms for the different forms of the test comparable. Accordingly, the Thorndike Mean would be 56.655, which is more than 8 greater than that of the above test. It was the intention to make the above test more difficult that the Thorndike test, and this has evidently been accomplished.

However, the test is by no means perfect. Sandiford, in his "Educational Psychology", pages 318 et seq. lists the criteria of a good test as validity, accuracy, reliability, objectivity, economy of time and effort and satisfactoriness of norms. Its validity is at the moment problematic; but we do know one thing , and that is that it does not measure the vocabulary the child uses of his own accord. Confirmation of this statement will appear later.

V. M. Sims made an investigation concerning the "Reliability and Validity of Four Types of Vocabulary Tests", which he discusses in the Journal of Educational Research, volume 20, page 94, 1929. He chose seventy words from the Thorndike Test of Word Knowledge, Form B, which he submitted to one hundred and ten students in Grades V to VIII. The words were arranged in four tests, Identification, matching, multiple-response (such as ours), and checking. He comes to the conclusion that the co-efficient of reliability of the Multiple-Response form is +0.84, somewhat lower than that obtained in this investigation, but still quite high.

Validity he measured by a series of correlations between the scores obtained on the different tests. His co-efficients of the different tests with the Multiple-Response test were as follows:

Identification	.74
Matching	.93
Multiple-Response	.84
Checking	.54

The final conclusion he reaches is that all forms of the test except the checking test have a high degree of validity. Although the matching test is the best, the Multiple-Response test is very high also in this respect.

Accuracy Sandiford defines as follows: "A test is perfectly accurate when the units of measurement are wholly appropriate and absolutely equal at all points of the scale". Inasmuch as the number of words chosen from each thousand is not in this test fixed,

the assumption is a priori that just to that extent the test is not accurate.

Reliability has already been discussed.

The test is perfectly objective as two examiners marking one copy cannot get different results.

This test is quite economical in cost, and also in time taken to mark it. Even in its present rather unwieldy form (mimeographed sheets) a copy can be marked in a couple of minutes.

As for the norms, these are yet to be derived.

As mentioned above, rareness of a word does not always correspond with its difficulty. Thus No. 7 'immediate' is a comparatively common word, falling in the second thousand, but it appears quite difficult for some students in the lower grades studied. Its high rating probably comes from the inflected form 'immediately', as these two are considered as one word.

No. 73, 'coyotes' seems much easier than would be suggested by its position in the thirteenth thousand, but this is probably due to local association. Very few of them are seen in the New England states nowadays, and practically none in Great Britain.

Another word with a high degree of rareness is 'chum', which occurs in the eleventh thousand. Nevertheless, none of the subjects seemed to have any difficulty with it. This word probably has a less evident provincialism about it, and is not so common in other parts of the world as here.

No. 97, 'quandry' is spelt wrong, but this apparently has not affected its useful ness.

No. 10, 'accompany' really has two possible synonyms. After the test was set up and distributed, it was found that 'companion' has official sanction as a verb. However, 'attend' seems still the best choice, and accordingly is considered the correct answer.

A slight degree of ambiguity exists in No. 3, as 'notice' is occasionally used synonymously with 'sight', e.g. "It came to my notice". However, the equivocalness here is practically negligible.

In marking this test, no allowance was made for guessing. Concerning this type of test, Sandiford, in his "Educational Psychology", gives the following formula for deriving the final score:

$$\text{Score} = \text{Rights} - \frac{\text{Wrongs}}{N - 1}$$

where N is the number of responses presented. But he goes on to say that "If four or more choices are allowed, chance plays a comparatively insignificant part in determining the score and usually need not be considered". It was further felt that the element of guessing minimized still more by the suggested incorrect choice, as mentioned above; and by normalizing the distribution. However, we shall see that the element of chance is not so negligible as Sandiford would have us believe.

Results were obtained from all schools indicated in Table I, with the exception of Wallace School. In addition, this test was given as part of the University of Alberta School of Education

survey of the King George School at Leduc, and these results have been incorporated into the data in attempting to establish grade and age norms. From the King George School tests were received from twenty-three pupils in Grade XI, fifteen in Grade X, twenty-six in Grade IX. In all, the test was administered to five hundred and fifty-nine students, of whom eleven were not considered in finding grade averages as they had either ommitted to indicate the grade or the grade shown did not agree with that indicated on the October composition or other tests. In finding averages and trends, the results from both the original and the revised forms of the test were used, since the revision was very slight. The test was administered in all cases with the exception of the school in Leduc during the month of February, 1935. At Leduc it was given the March following.

"DISTRIBUTION BY GRADES"

Table V shows the distribution of the scores by grades, together with the mean for each grade and other such measures. Q_1 and Q_3 as usual stand for the first and third quartile respectively. On the diagrams the Median is indicated by Q_2. The Semi-interquartile Range $(Q_3 - Q_1)/2$, is indicated by "s". Diagrams XI to XVI show the actual and theoretical distributions in Grades VII to XII respectively.

In considering the distribution of Grade VII on Diagram XI, the most striking feature is the way in which the scores are piled

TABLE V

DISTRIBUTION OF VOCABULARY SCORES BY GRADES

TOTAL NO. OF CASES = 548

Vocabulary Score	VII	VIII	IX	X	XI	XII
96 - 100						2
91 - 95			1	2	8	13
86 - 90			2	2	7	16
81 - 85			4	6	13	10
76 - 80	1		13	10	17	20
71 - 75	0	4	13	11	14	15
66 - 70	1	7	9	19	13	6
61 - 65	3	6	21	8	12	5
56 - 60	5	13	11	12	9	4
51 - 55	8	17	18	8	3	2
46 - 50	11	12	6	4	3	
41 - 45	10	8	7	1	2	
36 - 40	16	10	3	1		
31 - 35	15	2	1	1		
26 - 30	6	3	1			
21 - 25		1				
No. of cases	76	83	110	85	101	93
Mean	42.63	50.93	61.82	66.15	72.15	77.55
Standard Deviation	10.57	10.25	12.80	11.90	12.15	10.10
Median	40.50	51.62	61.90	66.97	73.04	78.63
Q_3	48.82	58.57	72.08	70.57	81.83	86.49
Q_1	33.00	42.98	52.56	57.60	63.44	72.08
s	8.66	7.795	9.76	6.49	9.20	7.70

DIAGRAM XI

DISTRIBUTION OF VOCABULARY SCORES IN GRADE VII

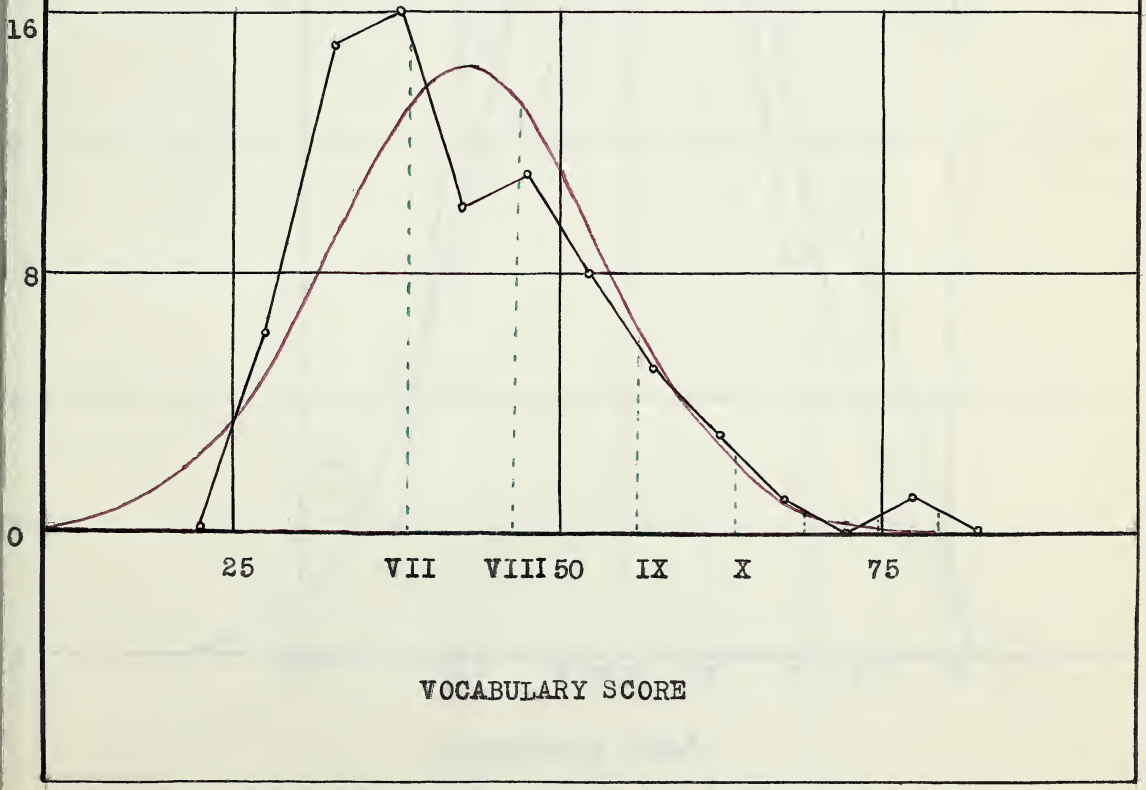
ACTUAL DISTRIBUTION ———

THEORETICAL DISTRIBUTION —

NO. OF CASES: 76

Mean	:	42.63
Sigma	:	10.57
Q_3	:	48.82
Q_2	:	40.50
Q_1	:	33.00
s	:	3.66

Successive Grade Levels shown by - - - - - - -



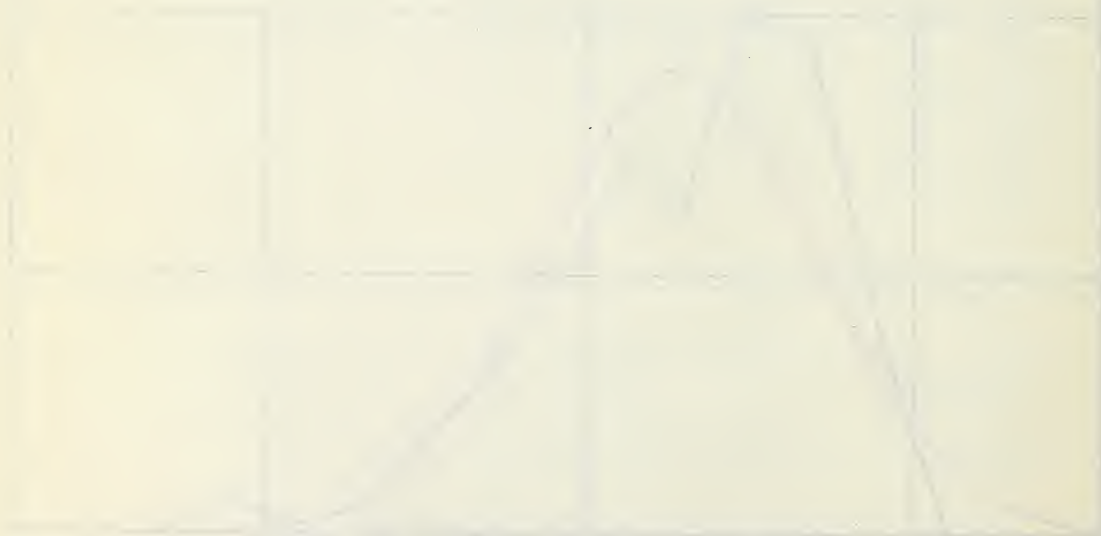


DIAGRAM XII

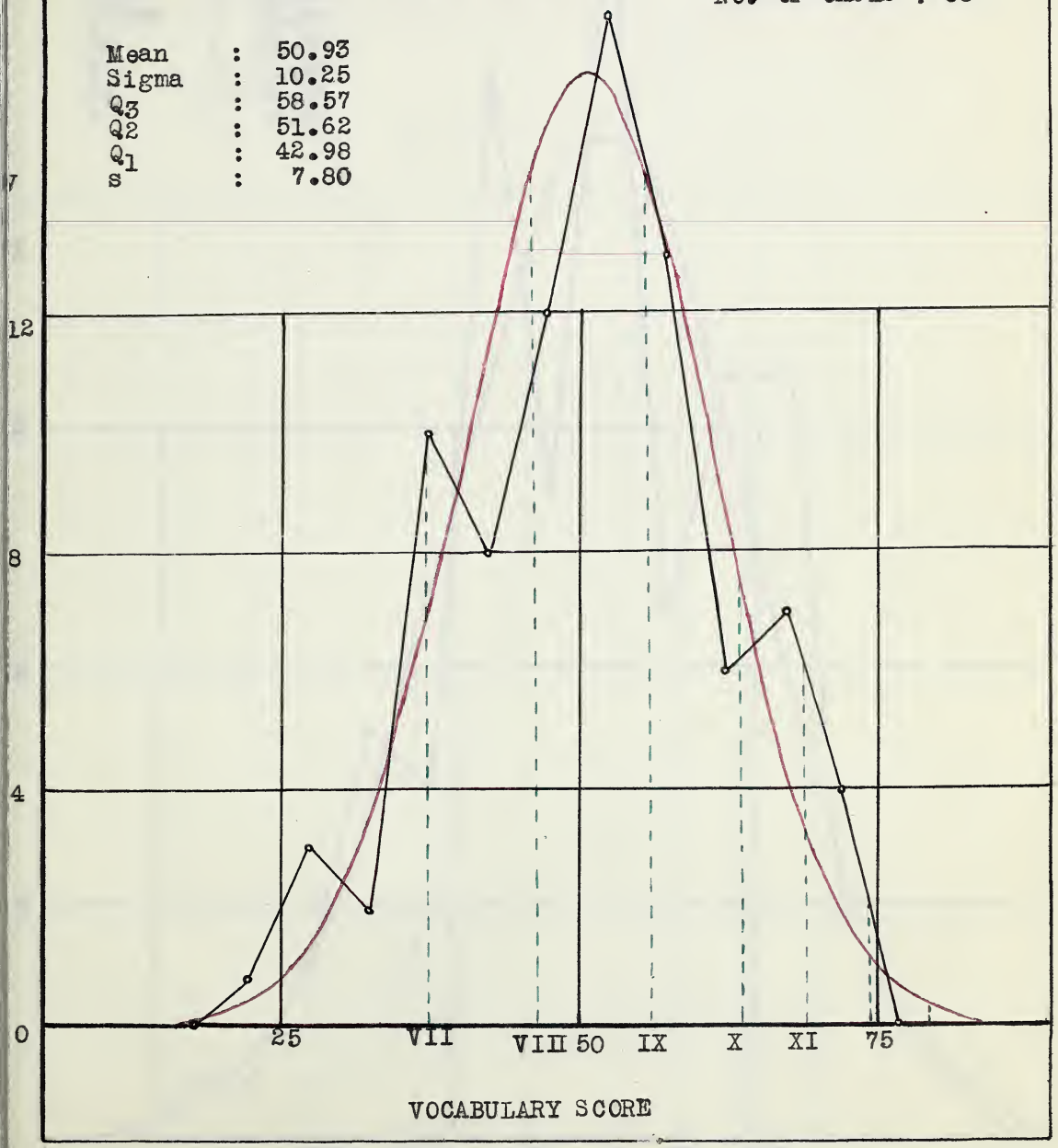
DISTRIBUTION OF VOCABULARY SCORES IN GRADE VIII

ACTUAL DISTRIBUTION ——— THEORETICAL DISTRIBUTION ———

SUCCESSIVE GRADE LEVELS SHOWN BY ———

NO. OF CASES : 83

Mean	:	50.93
Sigma	:	10.25
Q_3	:	58.57
Q_2	:	51.62
Q_1	:	42.98
s	:	7.80



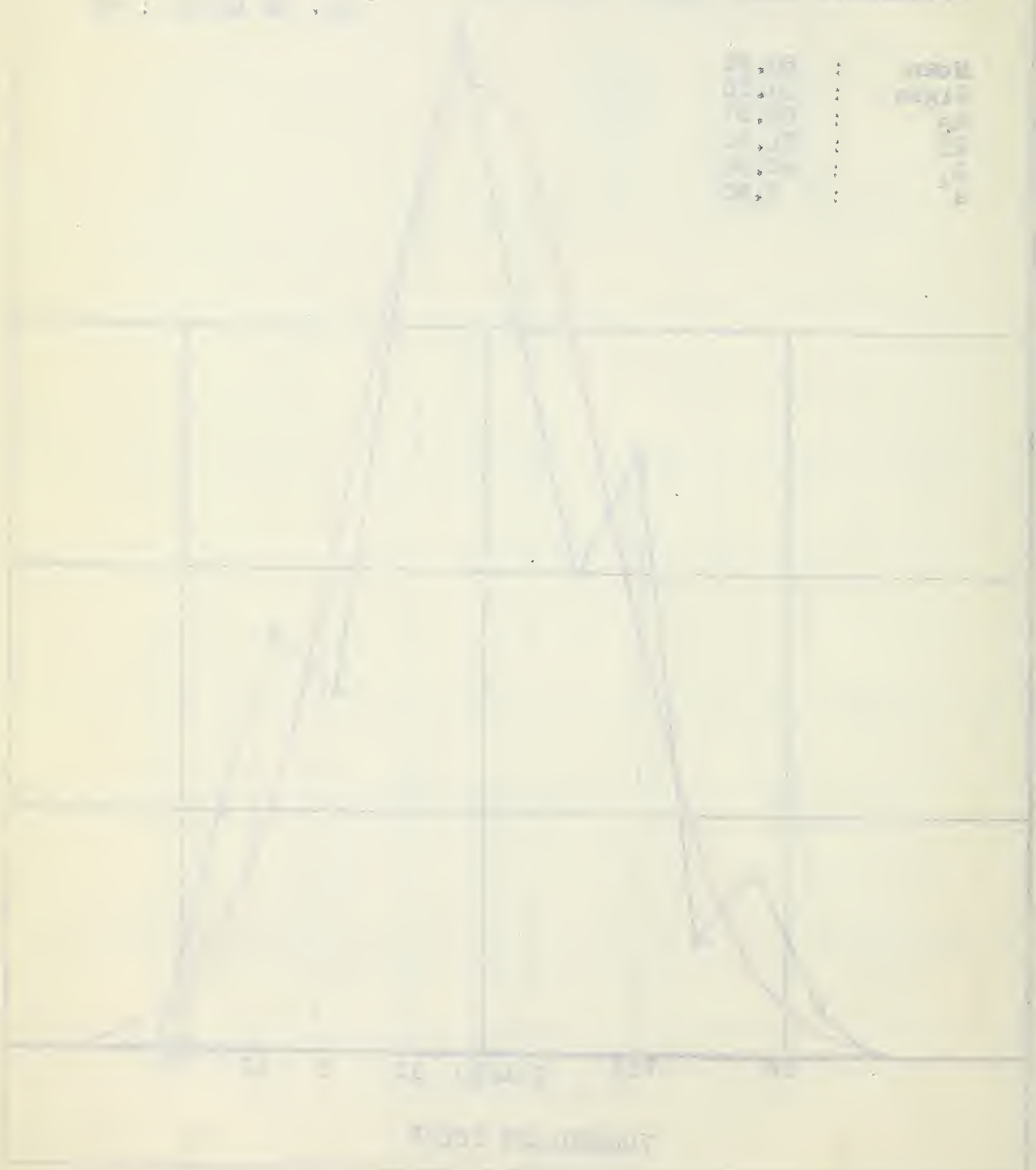


DIAGRAM XIII

DISTRIBUTION OF VOCABULARY SCORES IN GRADE IX

ACTUAL DISTRIBUTION ——— THEORETICAL DISTRIBUTION

SUCCESSIVE GRADE LEVELS INDICATED BY - - - - -

Mean	:	61.82
Sigma	:	12.80
Q_3	:	72.08
Q_2	:	61.90
Q_1	:	52.56
s	:	9.76

NO. OF CASES: 110

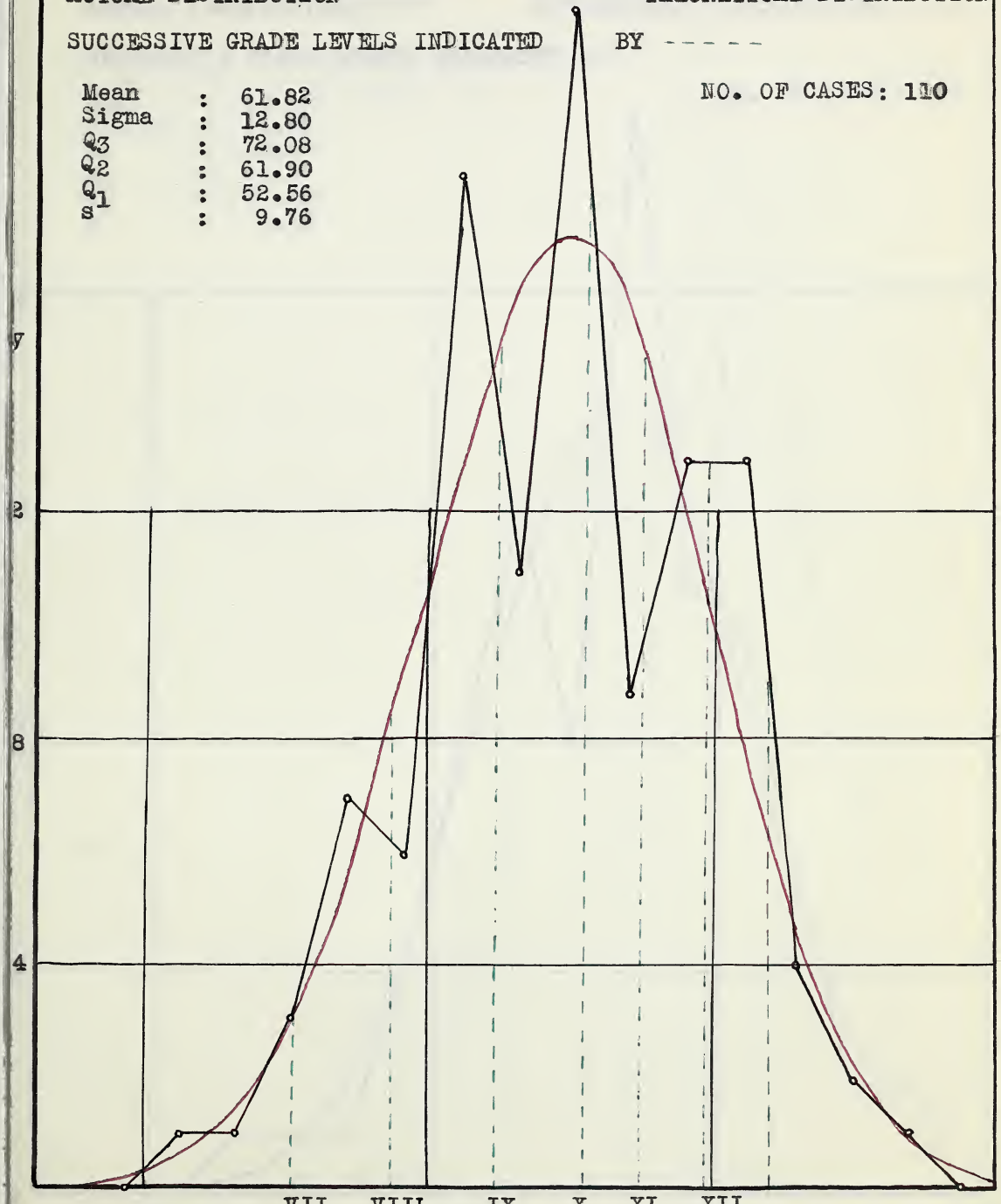


DIAGRAM XIV

DISTRIBUTION OF VOCABULARY SCORES IN GRADE X

ACTUAL DISTRIBUTION ——— THEORETICAL DISTRIBUTION———

SUCCESSIVE GRADE LEVELS INDICATED BY - - - - - -

NO. OF CASES: 85

Mean	:	66.15
Sigma	:	11.90
Q_3	:	70.57
Q_2	:	66.97
Q_1	:	57.60
s	:	6.49

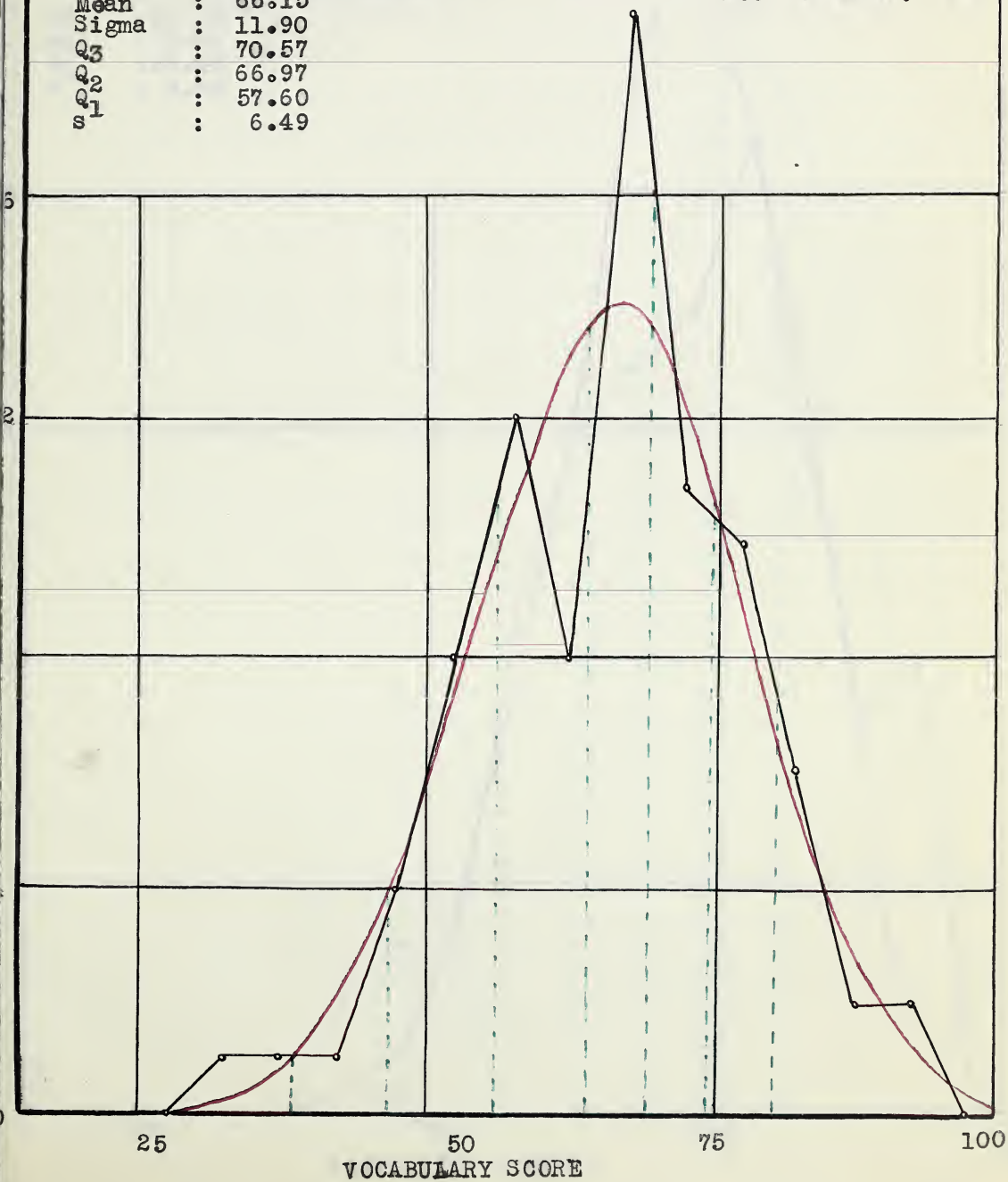


DIAGRAM XV

DISTRIBUTION OF VOCABULARY SCORES IN GRADE XI

ACTUAL DISTRIBUTION ——— THEORETICAL DISTRIBUTION———

Mean :72.15
Sigma :12.15
Q_3 :81.83
Q_2 :73.04
Q_1 :63.44
s : 9.20

SUCCESSIVE GRADE LEVELS INDICATED BY - - - - -

No. of Cases: 101

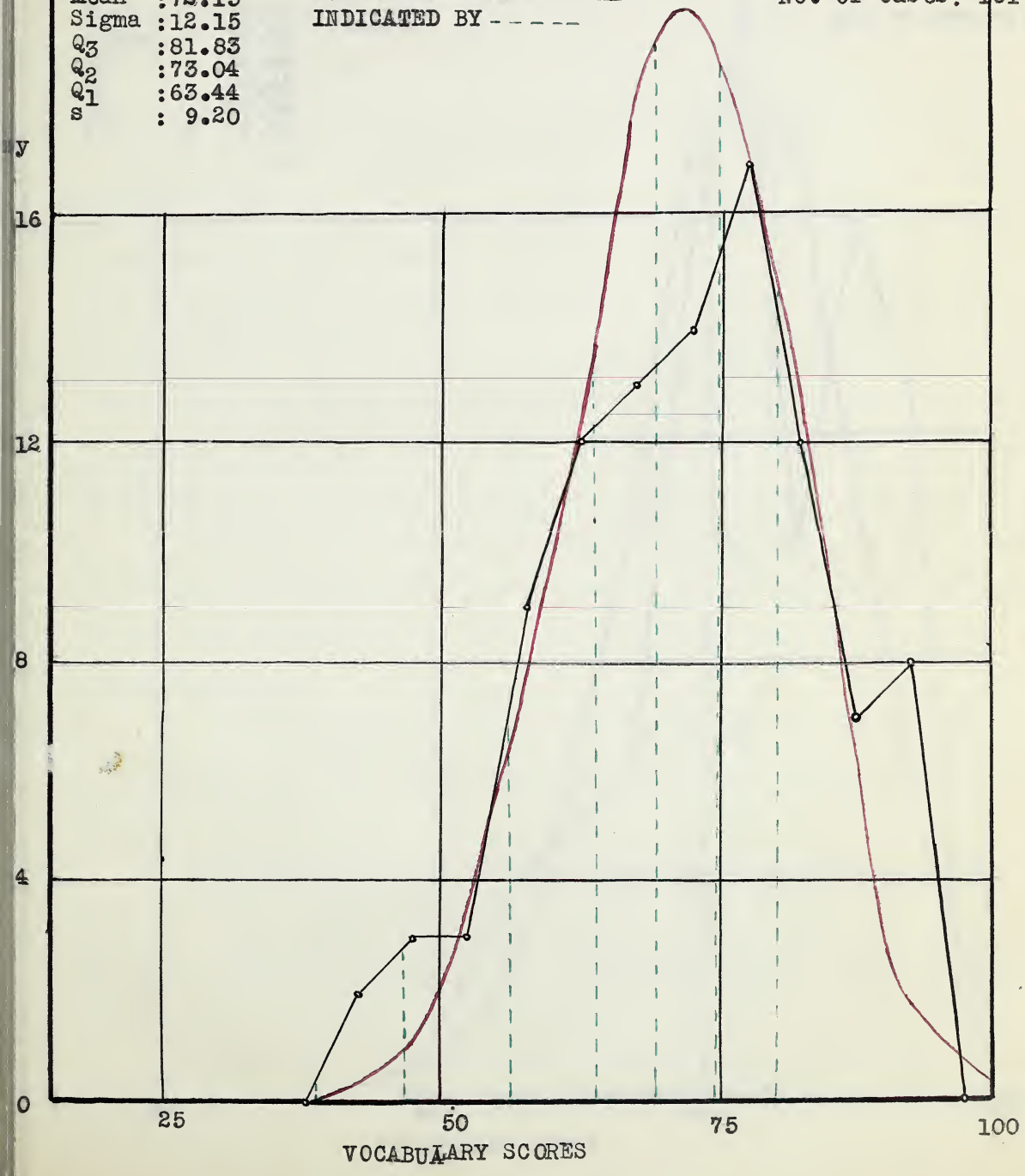


DIAGRAM XVI

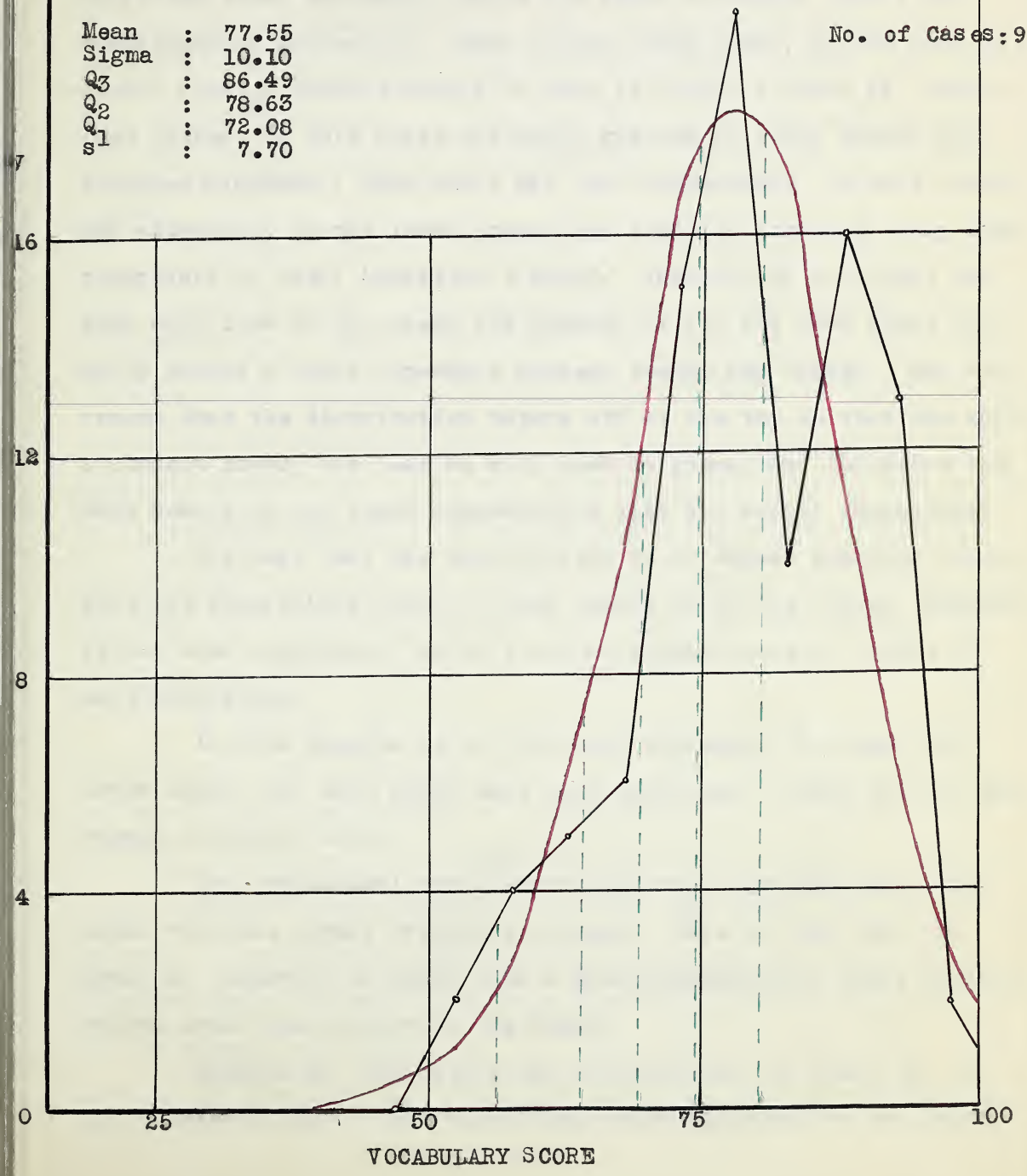
DISTRIBUTION OF VOCABULARY SCORES IN GRADE XII

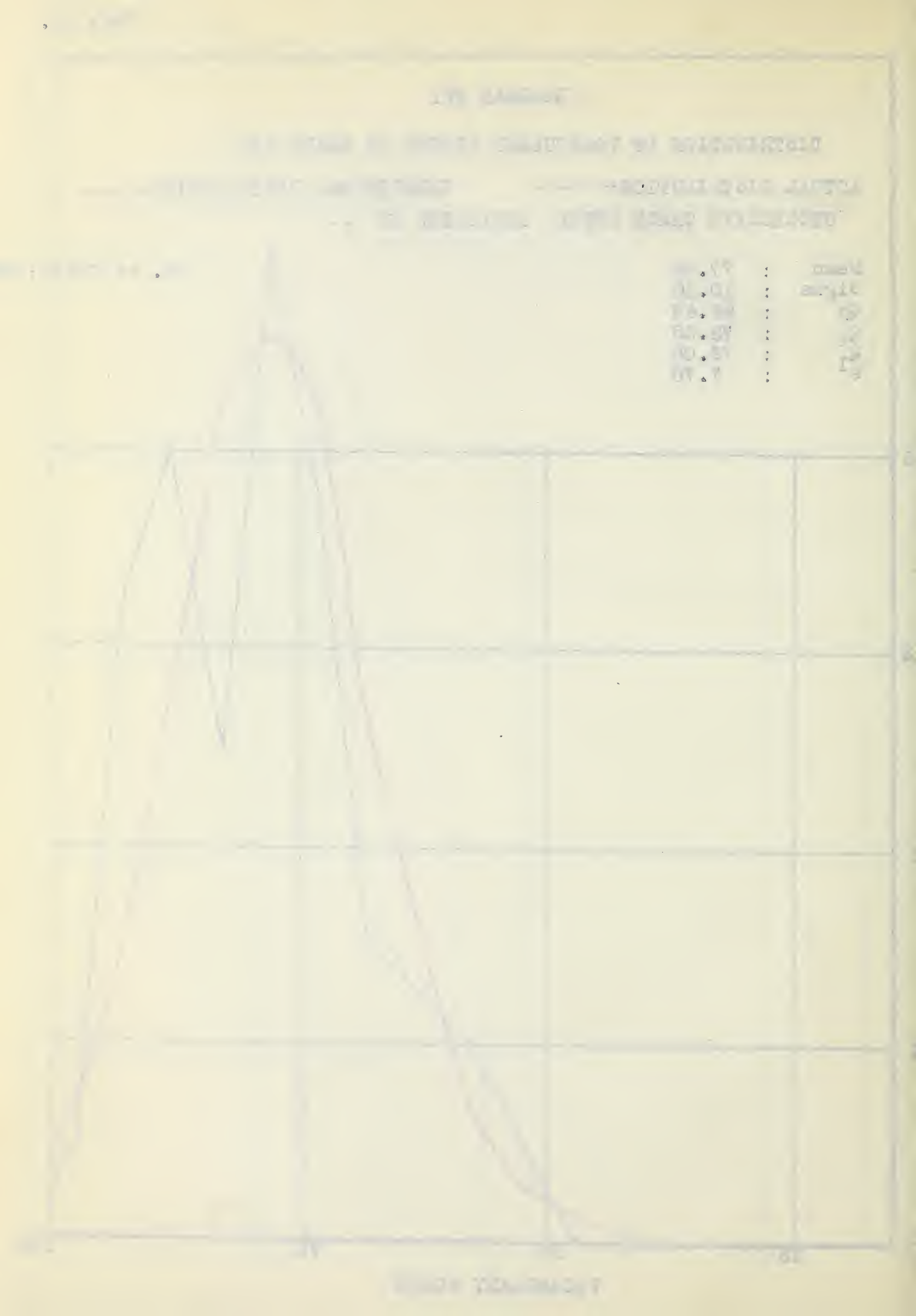
ACTUAL DISTRIBUTION——— THEORETICAL DISTRIBUTION———

SUCCESSIVE GRADE LEVELS INDICATED BY - - - - - -

Mean	:	77.55
Sigma	:	10.10
Q_3	:	86.49
Q_2	:	78.63
Q_1	:	72.08
s_1	:	7.70

No. of Cases: 93





up at the lower extremity, while the upper extremity tapers off comparatively gradually. Other things being equal, by the laws of chance alone, a student should be able to obtain a score of twenty. (One student in this grade obviously guessed at every answer and obtained nineteen . This score has been discarded). In most cases, and especially in the lower grades, the students answered every item regardless of their knowledge thereof. Accordingly a student who knew only five or ten items and guessed at all the rest would normally obtain a score somewhere between twenty and thirty. And the reason that the distribution tapers off at the top is that the more a student knows, the less he will need to guess, and therefore the more nearly is his score commensurate with his actual vocabulary.

The fact that the distribution is so skewed probably means that the theoretical curve is also pushed up on the scale. However, it has been considered "as is", and no attempt made to orient it more accurately.

In this diagram as in the ones following, the upper and lower limits for each grade have been indicated. These will be discussed in detail below.

The theoretical curve for Grade VIII only approximately coincides with the actual frequency polygon. This is also true for Grade X. However, no doubt with a greater population, these distributions would come closer to the ideal.

Diagram XI, indicating the distribution for Grade XI, is quite interesting in that the actual curve approximates the theor-

etical one very closely, except that frequencies which should be at the very centre of the curve appear to be pushed, not just a little to each side, but to the very extremities, upper and lower. About seventeen percent of the population tested are either quite inferior or quite superior to the milling herd as far as vocabulary goes.

The distributions for Grades IX and XII might provide very fine inspirations for Gothic architecture, but as approximations to the normal curve they are distinctly disappointing. Probably finer grouping would smooth them out somewhat, but would certainly not get rid of all the superflous modes. The explanation for these discrepancies probably lies in the fact that the populations in each grade are quite heterogeneous. In one respect at least they fall into two well-defined groups, unilingual and bilingual, and this no doubt has much to do with the peculiar distributions. The question of bilingualism will be fully dealt with later.

In Diagram XVII, the mean scores for each grade have been plotted, and the trend indicated. This trend line has merely been sketched in by sight. Since a child's grade is not a measure of ability or achievement in any single subject or groups of subjects, a more detailed study of trends has been left until ages are considered. A superficial examination of the trend line for grades, however, seems to indicate that it lies between a straight line and an exponential curve.

In Diagram XVIII the quartiles for the grades are plotted.

DIAGRAM XVII

GRADE MEANS FOR VOCABULARY TEST

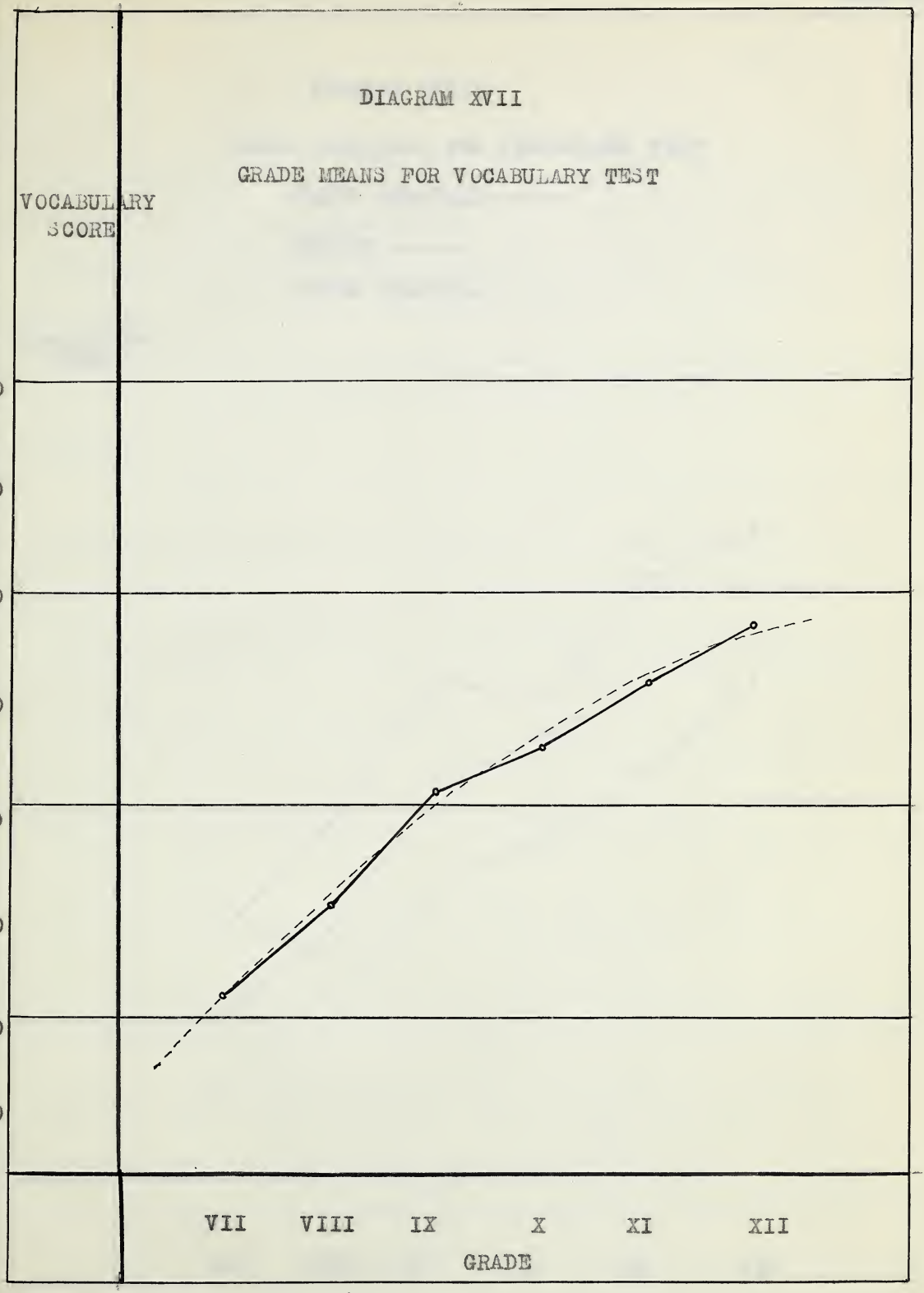


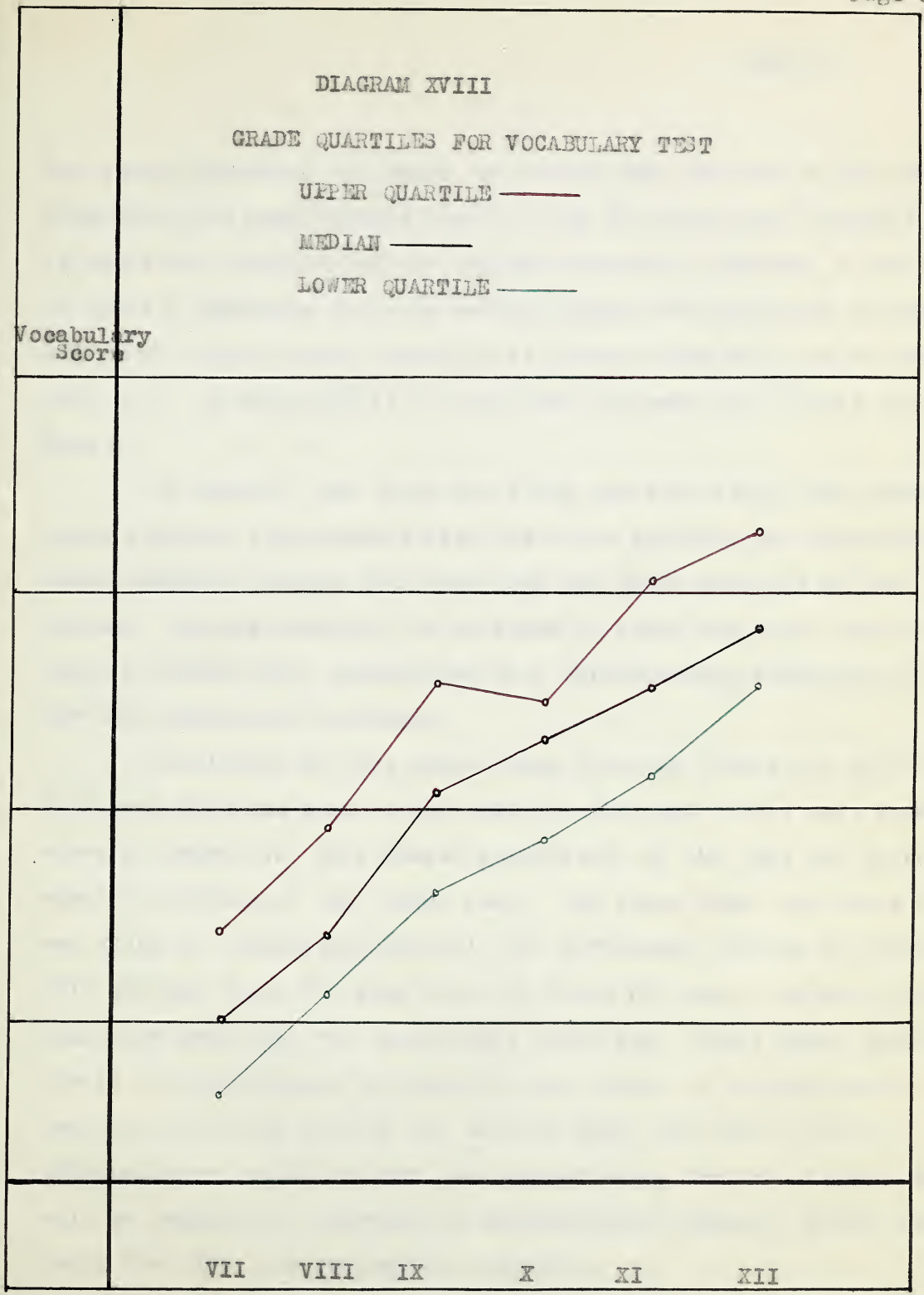
DIAGRAM XVIII

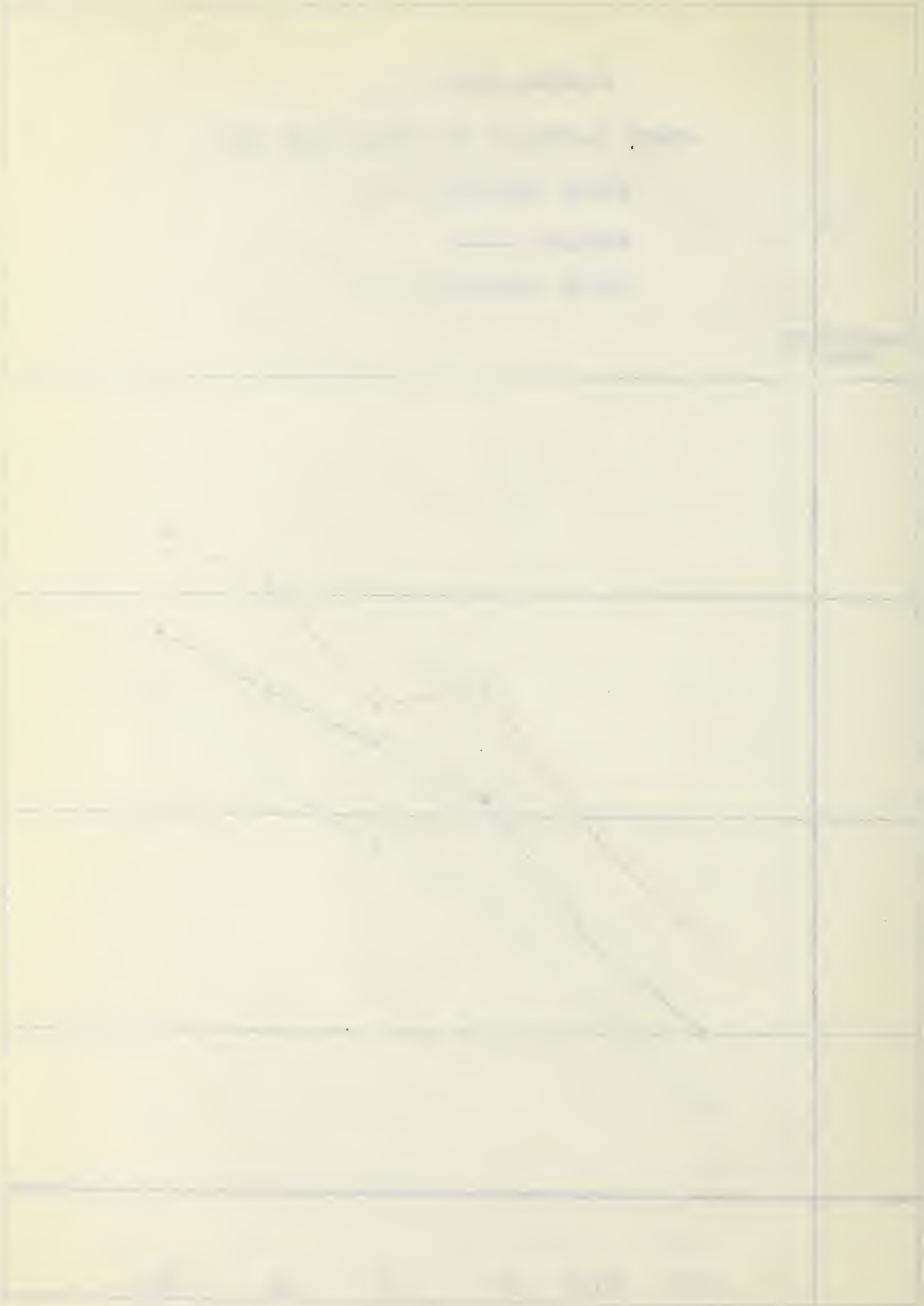
GRADE QUARTILES FOR VOCABULARY TEST

UPPER QUARTILE ——

MEDIAN ——

LOWER QUARTILE ——





The second quartile, or median would have been plotted on the same diagram as the mean, except that as they lie very close together, it would have complicated the diagram too much. However, a study of Table V indicates that the median follows the mean very closely and in all cases except Grade VII is greater than the mean by less than 1.1. In Grade VII it is less than the mean by a little more than 2.

In general, the first and third quartile follow the median quite closely, the distance from the third quartile to the median being somewhat greater than that from the first quartile to the median. The one exception is in Grade X, where the third quartile takes a sudden drop, accompanied by a corresponding diminution of the Semi-Interquartile Range.

The limits for the grades were found by taking one half the distances from the mean of one grade to the mean of the next grade above or below it. This seemed reasonable, as the test was given about the middle of the school year. The lower limit for Grade VII was found by subtracting one-half the difference between the Grade VIII and the Grade VII mean from the Grade VII mean, and the upper limit for Grade XII was established similarly. Since these grade limits are established by interpolation, which is only an approximation; or in the case of the extreme upper and lower limits by extrapolation, which is even less approximate, they can be considered only as tentative. However, an examination of Table VI would indicate that they are reasonably accurate.

It would perhaps have been statistically sounder to have

worked from the medians rather than the means. The reason for this is that in computating the mean one tacitly assumes that all the cases are concentrated at the centre of each group or class, whereas in the computation of the median, and other percentiles, the assumption is that the cases are distributed evenly throughout the class. And in finding the number of cases in each grade we have used exactly the same system as in finding percentiles. That is, we have subtracted the cumulative frequency to the lower grade limit from the cumulative frequency to the upper grade limit.

However, it was desired to find the theoretical as well as the actual number of cases at each grade level, using for this purpose the integrals under the normal curve. And since Camp's Table I (Elementary Statistics, p. 380) is based on the mean and derived product moments, we have chosen the mean as our grade norms.

The grade limits so derived are as follows:

Grade	Mean	Limit
XII		80.25
	77.55	
XI		74.85
	72.15	
X		69.15
	66.15	
IX		63.98
	61.82	
VIII		56.38
	50.93	
VII		46.78
	42.63	
		38.48

Table VI (a) indicates the percentage of students in each grade who are at each of the other and the same grade norms. Thus, of the pupils in Grade VII, 42.30% are below Grade VII standard,

TABLE VIa

OVERLAPPING OF VOCABULARY SCORES BY GRADES
ALL FREQUENCIES EXPRESSED AS PERCENTAGES.
A. AND T. STAND FOR ACTUAL AND THEORETICAL REL/
ATIVE FREQUENCIES

Grade	Below VII		VII		VIII		IX		X		XI		XII		ABOVE XII	
	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T
VII	42.30	34.5	23.18	31.0	21.70	25.0	7.90	6.2	1.89	1.8	.0016	.3	1.45	0.0	-	0.0
VIII	15.60	11.3	18.60	23.0	43.60	36.0	12.30	19.5	3.95	6.5	6.70	3.1	.0016	0.4	-	0.2
IX	3.72	3.4	9.140	8.5	23.50	21.5	22.45	23.3	10.68	14.9	12.83	13.0	12.34	7.8	6.18	7.6
X	2.00	.4	3.22	4.9	16.32	15.3	17.40	22.1	19.90	17.3	16.95	16.8	12.53	11.4	11.41	21.8
XI	-	.3	2.97	1.6	7.33	6.8	15.90	16.0	13.10	14.5	15.65	18.3	17.89	14.6	27.02	26.6
XII	-	0.0	-	.1	3.34	1.7	7.40	7.2	6.45	11.3	16.75	19.2	22.45	21.0	43.60	39.5

TABLE VIb

RETARDATION AND ACCELERATION OF VOCABULARY SCORES BY GRADES

ALL FREQUENCES EXPRESSED AS PERCENTAGES

A. AND T. AS FOR TABLE VIa

Grade	Below grade Norm		AT Grade Norm		Above Grade Norm	
	A	T	A	T	A	T
VII	42.30	34.5	23.18	31.0	34.52	34.5
VIII	34.20	34.3	43.60	36.0	22.20	29.7
IX	3720	33.4	22.42	23.3	46.38	43.3
X	39.20	42.7	19.90	17.3	40.90	40.0
XI	39.25	40.2	15.62	18.3	45.13	41.5
XII	34.14	39.5	22.24	21.0	43.60	39.5

21.70% at Grade VII, 7.90% at Grade IX, and so on. The theoretical as well as the actual percentages have been given. These have been derived from the applications of the normal curve as indicated in Diagrams XI to XVI respectively.

In these diagrams, the Grade limits as mentioned above, have been indicated. The area under the frequency polygon between the grade limits thus represents approximately the actual number of cases at each grade level; and the area between the same limits under the normal curve represents quite accurately the theoretical number at each grade level.

Occasionally we find both theoretical and actual frequencies at one grade level greater than at the next grade level towards the centre. Thus in Grade IX both the theoretical and actual frequencies of students at the Grade XI level are greater than the corresponding frequencies at the Grade X level. This is to be expected, as the widths of the grade levels are not uniform.

The number below Grade VII level decreases and the number above Grade XII level increases as we go from the lower to high grades. This is naturally so as the mean with each successive grade increases, or moves from left to right in our table.

Table VI (b) indicates the theoretical and actual relative frequencies of each grade below, above, and at that grade level. Generally speaking, about twenty percent of the population of each grade, is at its own grade level, with thirty-five to forty-five percent above and below. This would indicate that despite assumptions, extrapolations, and what-not, the grade limits have been

fairly well chosen. However, Grade VIII is an exception, with well over forty percent at its own grade level, and just over twenty percent above that grade level.

DISTRIBUTION BY AGES

In grouping the data according to the ages of the subjects, eleven more cases, as mentioned above, are considered than in the grouping according to grades. This gives us a total frequency of five hundred and fifty-nine, distributed as shown in Table VII.

The ages of the subjects were taken as at their last birthday, which might range anywhere from no days at all before the test was written to just short of a full year. It was thought inadvisable to attempt a grouping of less than a full year, as such groups would contain very few frequencies and be much more subject to the fluctuations of chance. Instead, all cases are considered as concentrated at the half-way point from one year to the next so that an age of sixteen is taken as being sixteen and one-half years. Norms or averages for finer divisions than one year can be found either by interpolating in Table VII or Diagram XIX, or by substituting the age in the regression equation.

Although all available cases have been considered, means and other measures of ages below twelve and above nineteen cannot be considered reliable, as the greatest frequency for any one of those groups is six. Accordingly, in Diagram XIX, the lines connecting the mean points to the rest of the curve are dotted.

TABLE VII

DISTRIBUTION OF VOCABULARY SCORES BY AGES

TOTAL NO. OF CASES = 559

Vocabulary Score	AGES 11	12	13	14	15	16	17	18	19	20	21	22	23	24
96 -100						1					1			
91 - 95					6	5	6	5	1				1	
86 - 90			1	1	3	3	11	3	4					
81 - 85			1	6	1	9	9	5	1		1			
76 - 80		2	2	9	10	14	9	9	3	1				
71 - 75			3	10	9	11	12	5	5	1	1			
66 - 70		3	6	7	14	15	5		1	2				
61 - 65		1	7	12	14	10	7	1	1			1		
56 - 60	1	5	10	7	13	8	6	4	2					1
51 - 55	2	7	12	11	13	10	2	1	2					
46 - 50	1	8	6	7	5	8	3	1						
41 - 45	1	2	7	12	1	5		2			1			
36 - 40	1	6	12	4	7	3	1							
31 - 35		5	5	4	5			1						
26 - 30		2	3	3	3	1								
21 - 25						1								
No. of Cases.	6	41	75	93	102	103	71	37	20	4	4	1	1	1
Means	48.35	48.35	31.55	58.20	61.85	65.60	73.70	73.20	73.50	71.25	74.00	62.50	92.50	57.50

Let V stand for Vocabulary Score, and A for age in years.

Mean of V = 66.28 Mean of A = 15.72
Standard Deviation of V = 17.58 S. D. of A = 2.15
coeffieient of correlation, r_1 = .215 $\pm$ .027
Regression coefficient of V on A m_a = 1.75; of A on V m_V = .0269
Therefore

$$V = 62.28 + 1.75 (A - 15.72)$$
$$A = 15.72 + .0269 (V - 62.68)$$

Correlation ratio, Eta of V on A = .47
Therefore

$$1 - (Eta)^2 = .779$$
$$(Eta)^2 - r^2 = -.24$$

DIAGRAM XIX

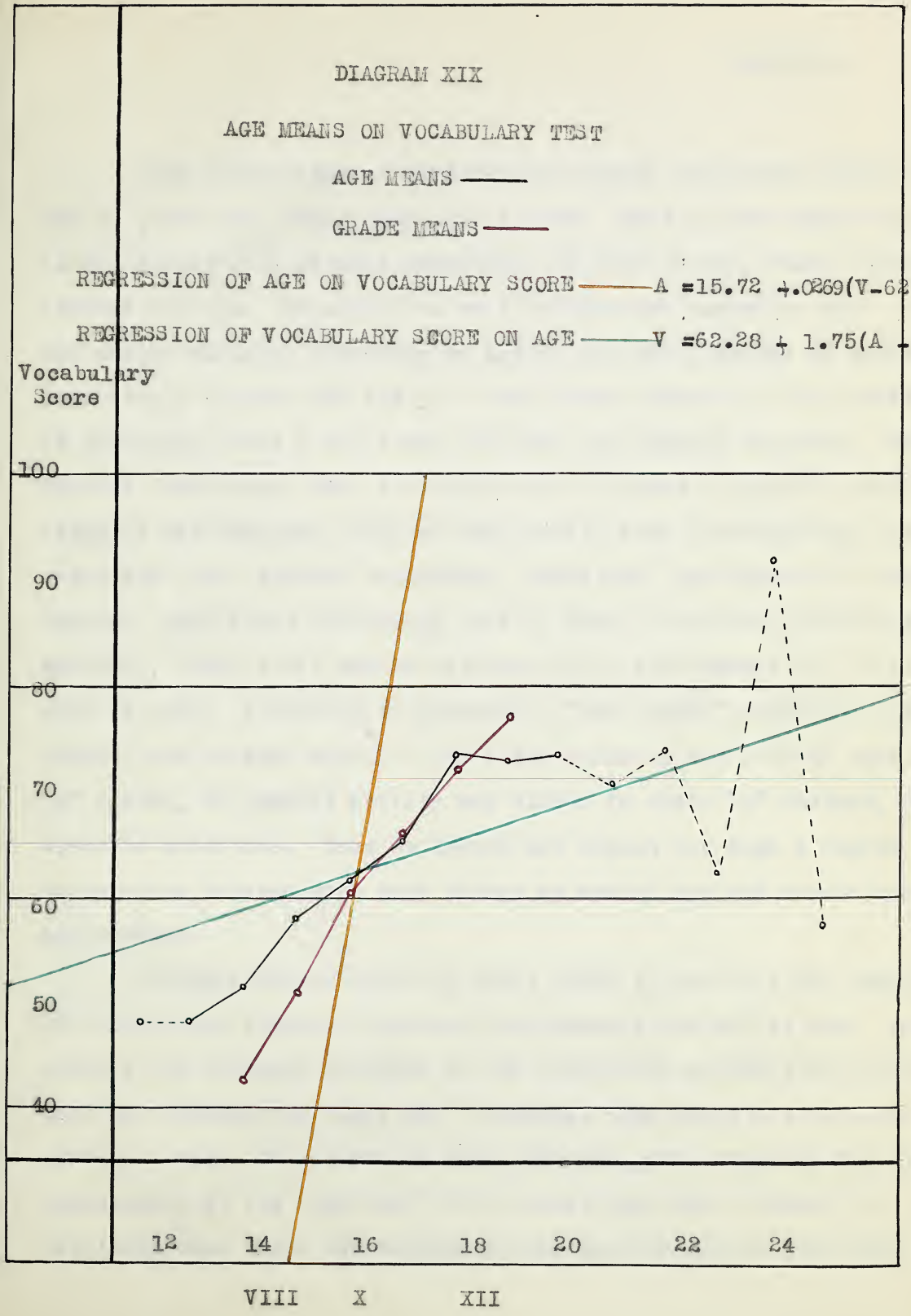
AGE MEANS ON VOCABULARY TEST

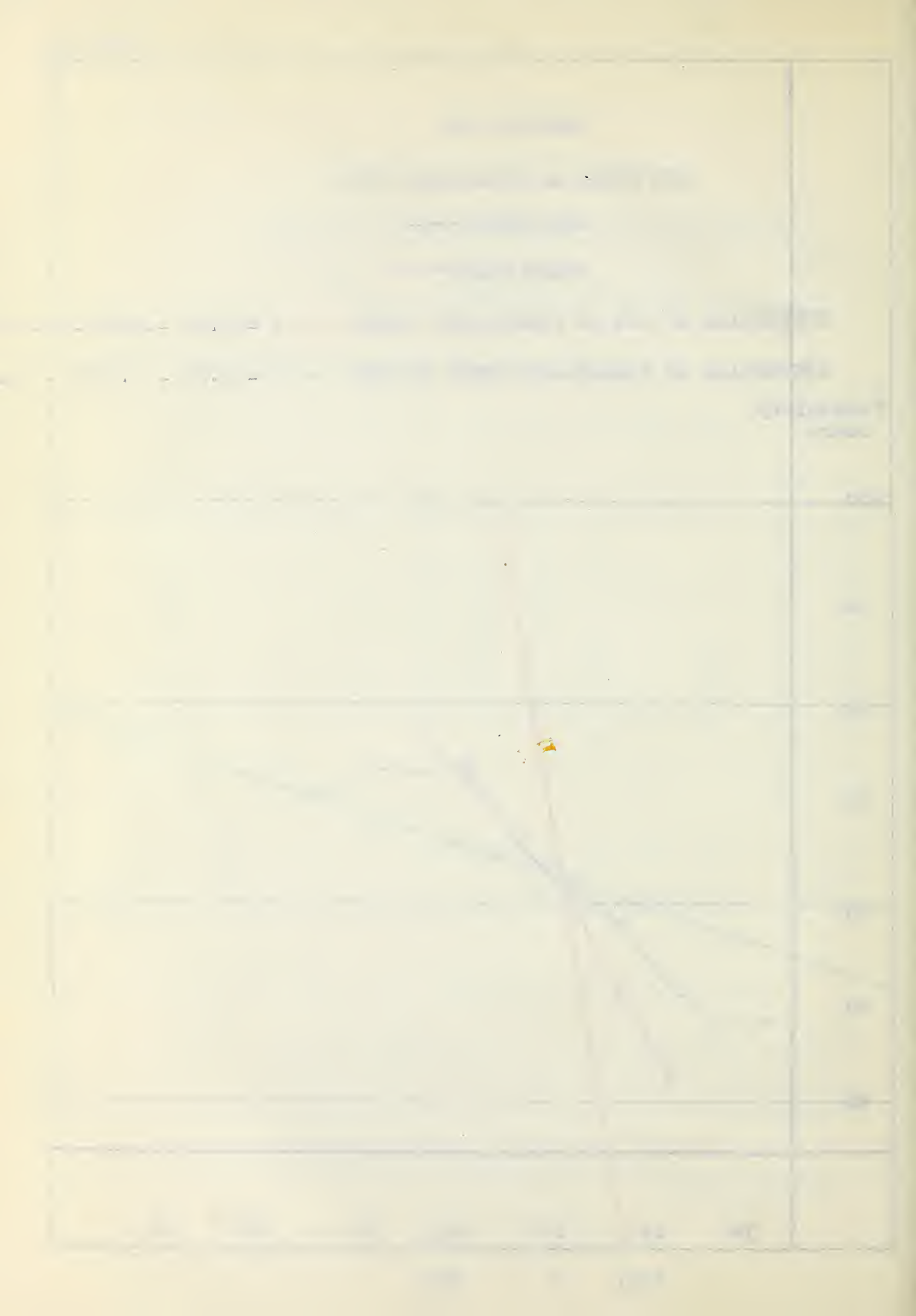
AGE MEANS ——

GRADE MEANS ——

REGRESSION OF AGE ON VOCABULARY SCORE —— A =15.72 +.0269(V-62.68)

REGRESSION OF VOCABULARY SCORE ON AGE —— V =62.28 + 1.75(A - 15.7





The coefficient of correlation between vocabulary score and age is quite low, being only .215 ± .027. This is not surprising, as the extent of a child's vocabulary depends on many other factors besides his age. Possibly, in fact it does not depend on age at all, but varies directly according to age to the small extent to which it does simply because his age is a very crude measure of the breadth of experience that a child has in using the English language. Many factors indubitably tend to affect and influence a person's knowledge of the language, such as familiarity with the vernacular in every day life, cultural background, education, and degree of intelligence. Even where everything else is equal, including Intelligence Quotient, there still may be differences in achievement on a test such as ours. According to Spearman's "two factor" theory of intelligence, two persons equal in all other respects and with an equal "g" factor, or general ability may differ in their "s" factors, or specific abilities. Thus we should not expect too high a degree of correlation between even such things as mental age and vocabulary achievement.

Another factor which no doubt tends to minimize the degree of correlation between vocabulary achievement and age is that presumably the youngest children in our population are more brilliant than the average for their age, otherwise they would not be accelerated a grade or two and so come into the grade range of our study. Conversely, at the upper end of the scale the older students are less brilliant than their contemporaries who have graduated from high

school a year or so ago. Theoretically these two factors should cancel each other, but practical experience and surveys such as those conducted by the Alberta School of Education in Fort Saskatchewan, Stoney Plain, Bruderheim, and Lamont show us that it is much easier to keep a student in a grade longer than one year than to promote a child ahead of his classmates; especially as departmental regulations state a minimum but no maximum time that must be spent in certain grades. The net result, therefore, is that as a rule the children of less than average ability are retarded much more than are the superior students accelerated. And accordingly the correlation between age and vocabulary score is less than what a random sampling of the population at the upper age limits would give us. However, as inferred above, even under the best of conditions, we should not expect such relationship to be very marked.

Letting V stand for vocabulary score, and A for age, the following measures were derived from Table VII.

	V	A
Mean	66.28	15.72
Standard Deviation	17.58	2.15
Coefficient of Correlation	$r = .215 \pm .027$	
Regression coefficient	$m_a = 1.75$	$m_v = .0269$

$$V = 62.28 + 1.75\ (A - 15.72) \qquad A = 15.72 + .0269\ (V - 62.68)$$

Correlation ratio, Eta of V on A = .47

$$1 - (Eta)^2 = .779$$

$$(Eta)^2 - r^2 = .17$$

The first regression equation tells us what vocabulary score should be achieved at any given age; the second at what age

standard any given score represents. Analogously with the use of "M.A." for mental age and "C.A." for chronological age, we might say it gives us the V.A., or vocabulary age. Both the regression lines corresponding to these equations have been indicated on Diagram XIX.

In this study only the correlation ratio of V. on A. has been considered, because extent of vocabulary is, to some small extent at least, a function of age, but the reverse is not true.

The correlation ratio, Eta, is a measure of dispersion of column means and the more closely Eta approaches 1, the less is the dispersion. Accordingly, the smaller the value of $1 - (Eta)^2$, ranging from 0 to 1, the less is the dispersion. But for this expression we have a value of .779, which is quite large indeed.

The smallness of the abbolute value of $(Eta)^2 - r^2$, also varying between the limits of 0 and 1, gives us a measure of the linearity of regression, or the closeness with which the age means approximate a straight line. This measure for our population is .17, with a Probable Error of .023, as derived from the formula $.6745 \times 2\sqrt{(Eta^2 - r2)/N}$, where N is the number of cases, in this case 559. This definitely indicates non-linearity, which is corroborated by an inspection of Diagram XIX. Vocabulary achievement, therefore, is not, mathematically, a simple function of age.

On the same diagram are plotted grade means. Assuming that a child progresses at the rate of 1 grade per year, at the age of thirteen and one half, he would be half way through Grade VII, and

so on. The upper three grade means approximate the corresponding age means quite closely, but in the lower grades the corresponding age means are much higher. Probably this is true because whatever acceleration does take place occurs in the primary and intermediate grades of the public school and practically none after Grade VII is reached, It is here that the departmental regulations requiring a minimum time in each grade are introduced. Accordingly the more brilliant students are affecting the grade mean of the grade or grades above that corresponding to their own age.

Because of the great number of factors affecting a child's vocabulary, age norms alone cannot be considered satisfactory, measuring as they do only the length of time which he has had to acquire a word knowledge. Furthermore, they appear to be unduly influenced by the less clever students, or at least the students of less vocabulary attainments than their contemporaries. On the other hand, students in the same school grade are, supposedly at least, at the same level of academic attainment, although in this attainment a knowledge of English is only one factor among many. It is probably best to use both and let individual circumstances guide one.

CHAPTER VI

RELATIONSHIP BETWEEN VOCABULARY AND OCCUPATIONAL STATUS.

Up to this point we have made several assumptions, such as those concerning the effect of intelligence, social status, and bilingu alism, on a child's command of the English language, for the purpose of explaining various results of our data. It would have been incorrect to make those assumptions without some evidence to support them, and that is the purpose of this chapter and the two immediately following.

The present chapter will be devoted to discovering whether there is any relationship between the score a child obtains on the vocabulary test and the occupation of his father, or occasionally of his mother. All those who wrote the test were asked to write down the occupation of their respective fathers, and these have been classified into seven groups. The classification is by no means precise, as often an occupation would seem capable of being classified in two or more groups. The classifications, together with the vocations included therein are as follows:

1. Agricul tural: farmer
2. Commercial: Commercial traveller, merchant
3. Managerial: credit manager, banker, life insurance inspector, mine manager, lumber yard manager, dairy inspector, railroad conductor, grain company superintend-

ent, contractor, hotel manager, creamery manager, mine official, wholesale warehouse manager, supervisor of city post office, telephone exchange manager, secretary-treasurer of a refining company.

4. Clerical: ticket agent, mine storekeeper, municipal secretary-treasurer, civil servant, mail clerk, railway clerk, court reporter, notary public, police desk sergeant, bookkeeper, clerk.

5. Professional: research professor, doctor, dentist, clergyman, auditor, accountant, mining engineer, road engineer, civil engineer, optometrist, missionary, lawyer, teacher, druggist.

6. Skilled: mechanic, electrician, butcher, painter, glass blower, glass worker, shoemaker, baker, blacksmith, photographer, tanner, carpenter, telegrapher, tailor, plumber, printer, steam engineer, locomotive engineer, assistant locomotive foreman, boiler engineer, brick layer, brick burner, fire boss, railway hostler, pit boss, mine tipple boss, telephone wire chief, moulder, barber, harness maker, tinsmith, cook.

7. Unskilled: janitor, drayman, miner, bill poster, railway section laborer, nigh watchman, gardener, factory employee.

In addition to these occupations, there are a number of others that did not fit readily into the above scheme, and to avoid making it too cumbersome they were omitted from the classification.

Some of these occupations were as follows: lumberman, broker, cattleman, elevator-man, pumpman, private investigator, collector, letter carrier, oil agent, pool hall proprietor, musician, truck and taxi driver, bank messenger, and so on. A few others, such as engineer, agent, foreman were too vague for one to know approximately to what group they belong. Some students answered the question as to their father's occupation by "Unemployed", "Retired", or merely left a blank. There were accordingly a large number of students not included in this phase of the study. However, in all, three hundred and fifty-one subjects were classified.

In order to iron out grade differences, the vocabulary scores were tabulated according to their distance from the grade mean measured in terms of the grade sigma or Standard Deviation. Then the frequencies in each occupational division were collected. Thus, considering the mean as the origin, and sigma as the unit of measurement, all scores regardless of grade, lying between the origin and + 1 sigma were grouped together for each vocation and so on. This gives us Table VIII where A is the actual frequency and % the percentage in each vocational group.

In Grade XI, the origin minus sigma happened to come to exactly 60.0, and accordingly a score of sixty is just on the dividing line between the two groups— origin minus sigma, and origin minus 2 sigma. When this happened, the frequency was split, half being assigned to the lower and half to the upper group. This is why half frequencies occur in Table VIII.

In a perfectly normal distribution chosen entirely at random,

TABLE VIII

DISTRIBUTION OF VOCABULARY SCORES ACCORDING TO OCCUPATIONAL STATUS.

TOTAL NO. OF CASES = 351

ORIGIN AS GRADE MEAN

Deviation	Commercial		Clerical		Agricultural		Managerial		Professional		Skilled		Unskilled	
	A	%	A	%	A	%	A	%	A	%	A	%	A	%
over 3σ									1	2.2				
2σ — 3σ					1	.9	1	1.9	6	13.0	3	3.94	1	1.0
σ — 2σ	10	22.8	6	21.4	17	15.7	14	26.9	16	34.8	9	11.8	12	12.1
o — σ	10	22.8	11	39.3	26	24.0	23	46.2	12	26.1	34	44.7	37	37.4
o — —σ	20	45.6	10	35.7	33.5	31.0	6.5	12.5	10	21.8	21.5	28.3	39	39.4
—σ — —2σ	4	9.1	1	3.6	21.5	19.9	5.5	10.9	1	2.2	8.5	11.2	8	8.1
—2σ — —3σ					7	6.5	2	3.9					2	2.0
under — 3σ					1	.9								
TOTALS	44		28		107		52		46		76		99	

thirty-four percent of the total frequency will lie between the mean and plus or minus sigma, fourteen percent between plus and minus sigma and plus or minus 2 sigma respectively, and two percent between plus or minus two sigma and plus or minus 3 sigma. Thus the entire frequency lies between minus 3 sigma and plus 3 sigma. This provides us with a method for finding what we may call a coefficient of normality, and will tell us to what extent our sample is above or below an ideal universe from which we may suppose it to be drawn. From Table VIII frequencies below the origin have been, for computation purposes, considered negative and those above the origin positive. The percentage that each frequency is of the total of the vocational group in which it lies has been computed and considered positive or negative as the case may be, as have the ideal percentages mentioned above. Then the ideal percentages have been subtracted from the actual ones with due attention to signs, and the differences collected. An example will perhaps make this clearer. For instance, for the managerial group, the problem is as follows.

-3σ	-2σ	$-\sigma$	0	σ	2σ	3σ
−3.9	−10.6	−12.5	46.2	26.9	1.9	
−2	−14	−34	34	14	2	
subtract −1.9		3.6	21.5	12.2	12.9	−.1

The total of the differences is +50.2. Suppose we now divide by one hundred to make our final index more analogous in size with other co-efficients used in statistics. Taking it to two decimal places, we have a final coefficient of normality of + .50.

What are the limits between which this coefficient moves? Suppose the total frequency of our sample is farther from the mean

or origin of the ideal universe than either plus or minus 3 sigma. Then we have a problem like this:

	-3σ	-2σ	$-\sigma$	o	σ	2σ	3σ
—100%							
	—2	— 14	— 34	34	14	2	
sub-tract	—100	2	14	34	— 34	—14	—2

and the sum of the differences is —100. The coefficient of normality will therefore be —100/100 = — 1.0. The limits between which our index can move are therefore ± 1.0, and if the distribution is perfectly normal, the coefficient will equal 0.0.

What does the coefficient measure? It tells us to what extent our sample is above or below what it should be if drawn at random from a perfect universe. In this case it tells us to what extent the scores of the students in each occupational group are superior or inferior to those of the whole population studied.

Apologies are due for this digression into the realm of more or less pure statistics, but as far as the writer knows, this coefficient of normality is original with him, and therefore an explanation of it was in order.

The coefficients of normality for the various vocational groups were found to be as follows: Commercial —.21; Clerical .35; Agricultural —.25; Managerial .50; Professional .62; Skilled .22; Unskilled .01. Since over two hundred subjects are not represented, there is no way in which we can check our results. Theoretically, the sum of the products of our coefficient and the number of cases in each group would be zero.

It is rather surprising to find that the children in the commercial group are inferior to the population as a whole for usually merchants and salesmen are well up in the social and economic scale. However, with our comparatively small samples, extraneous factors such as lack of mental ability might easily enter, and we have no knowledge that our business men are representative of business men as a whole.

The students whose parents are in clerical occupations do better than the average. Clerical occupations, although usually not highly remunerative, are generally economically secure, and perhaps this is what we find here.

We might be tempted to ascribed the lowness of the coefficient for agriculture to the fact that thirty-four out of one hundred and seven students speak a foreign language; but the proportion for unskilled labor is even higher, and yet the children in this classification do much better than the farming population. Probably the truth lies in the fact that farmers notoriously have little leisure, and the same applies to their families. When a lad walks a mile or so to school and then home again, and has chores to do both before and after school, he is not going to have much time for reading. And his parents being constantly busy, they are not going to spend much money on books they wont have time to use.

Those students in the managerial group are much higher than any of the others, except the professional group. People in these occupations are generally well paid with a good deal of leisure and the education to use it.

As was expected the professional group ranks highest of any. This is no doubt partly due to the fact that students in this group live in homes where books are available; and probably these students inherit some of the intelligence and special aptitudes which made possible a professional education for their parents.

Those in the skilled group are somewhat above the average, which is not suprising as skilled craftsmen are generally intelligent and able men, who, in normal times at least, earn a good living. It is, however, rather surprising that the unskilled group do as well as they do, for generally these people are at the bottom of the social and economic scale, or at least near it. Furthermore, thirty-two of these come from homes where another language is spoken. Perhaps a more detailed study with a greater number of subjects would change this coefficient.

In conclusion, we may say that there is a definite relationshop between the vocabulary of a child and the occupation of his parent. In this study, with a comparatively small population and our classification made more or less arbitrarily on very scant data, we could not expect more definite results. It would be interesting to check the achievements in English against more complete data concerning the home environment, such as that furnished by Sims Socio-Economic Score Card.

No attempt has been made to allow for high or low intelligence in this chapter, because the test given, the Otis-Self-Administering, is entirely a reading test, and it was felt that a low intelligence score might arise from the same cause as a low vocabulary score, i.e., lack of ability in reading.

CHAPTER VII

EFFECT OF BILINGUALISM ON VOCABULARY

One hundred and nine of the students tested came from homes where another language besides English is spoken, and it was felt that a priori this would militate against them doing as well on the vocabulary as their classmates. Accordingly each student was asked to indicate what language or languages are spoken in his home. These languages were classified into four groups, as follows:

1. Central European: Ukrainian, Slav, Slovak, Polish Bohemian
2. Teutonic: German and Jewish
3. Scandinavian: Finnish, Swedish, Norwegian.
4. Romantic: French and Italian

It was rather surprising to find that of all the students tested, only two came from French-speaking homes. Apparently all the French-speaking students go to separate schools, which were not included in this study. However, since there are so many people in this country whose native tongue in French, no statement on bilingualism is complete that does not include a study of French-speaking subjects.

The scores were grouped according to the same method as in the previous chapter. That is, the distance from the grade mean, measured in terms of the Standard Deviation of the grade, was taken,

and then all the scores grouped according to languages. This gave a distribution as follows:

Deviation	Cen. Europe	Scandinavian	Teutonic	Romantic	TOTAL
3σ					
2σ	1	1	1	2	8
σ	12	4	8	5	29
o	13	4	11	6	41
−σ	13	1	7	2	23
−2σ	6		1		7
−3σ	1				1
TOTALS	49	10	28	15	109

These frequencies were reduced to percentages and the coefficients of normality found as in the previous chapter. They were as follows: Central European —.33; Scandinavian 0.0; Teutonic —.35 Romantic — .05. For the whole group it was —.33.

These indices are not ambiguous. In not one case is the coefficient positive. In other words, not one group does on the average better than the population from which they are selected; and only one, the Scandinavian, does as well. That, incidentally, being the smallest group is the one that gives us the least reliable coefficient. A larger sampling might prove it to be definitely inferior to the general population. And since these groups are below the average of which they themselves formed a part, they are even farther below their classmates who speak only English.

The literature on this, as on other phases of vocabulary study, is meagre; nevertheless our results corroborate those of previous investigators. In the Journal of Educational Psychology, vol. 24, 1933, page 696, there is an article entitled "The Essay Versus the Objective Examination As Measures of the Achievement of Bi-lingual Children", by Floyd F. Caldwell and Mary D. Moury. Six hundred and forty-three children in Grades III to VIII inclusive were tested and divided into two groups, Spanish-speaking and English-speaking. The children were tested in the fields of English and History by means of both an essay and an objective examination. Concerning the comparative achievements in the field of English, the results are summarized as follows: "The median scores of the Anglo-American children on objective and essay tests in English exceed those of the Spanish-American children by 18.1 percent and 28.6 percent, respectively,....."

Another recent study in this field is reported in the Journal of Educational Research, vol 27, 1934, under the title of "The English Handicap of Junior High-School Pupils from Foreign Speaking Homes, and Remedial Measures", by Ralph A. Fritz and Nellie R. Rankin. Two hundred and one students of Arma, Kansas, were tested, the tests comprising the Otis Self-Administering Mental Test, The New Stanford Achievement Test, and the Sim's Socio-Economic Score Card. The first method was to compare the two groups, English and foreign speaking; the second to compare pairs matched for intelligence of social and economic environment. Part of their results are

given below. In the column of differences, a negative sign indicates that the mean of the foreign group is larger; a positive sign the reverse. The test is the Stanford Achievement.

	Grade VII			Grade VIII			Grade IX		
	Eng.	For.	Diff.	Eng.	For.	Diff.	Eng.	For.	Diff.
No. of cases	19	13		40	10		28	11	
Test									
Paragraph Meaning	77.5	75.5	2.0	95.1	81.5	13.6	102.2	90.2	12.
Word Meaning	73.1	71.6	1.5	86.9	73.0	13.9	94.5	83.4	11.1
Dictation	82.5	78.1	4.4	91.2	83.0	8.2	101.3	92.9	8.4
Language Usage	69.9	52.8	17.1	90.1	71.5	18.6	96.6	77.9	18.7
Literature.	56.0	66.3	-10.3	72.1	55.5	16.6	82.9	78.4	4.5
SUM			15.7			70.9			54.7

Both by the method of comparing whole groups and the method of matched pairs these investigators come to the conclusion that foreign children suffer a handicap in school work, and that this handicap is greater in English than in other fields. A plan for remedial teaching is evolved, with particular stress on oral and written composition. No doubt such a policy would be very sound in Alberta, where apparently the foreign speaking children are also deficient in English, and where there are so many of them, almost twenty per cent of the population studied.

CHAPTER VIII

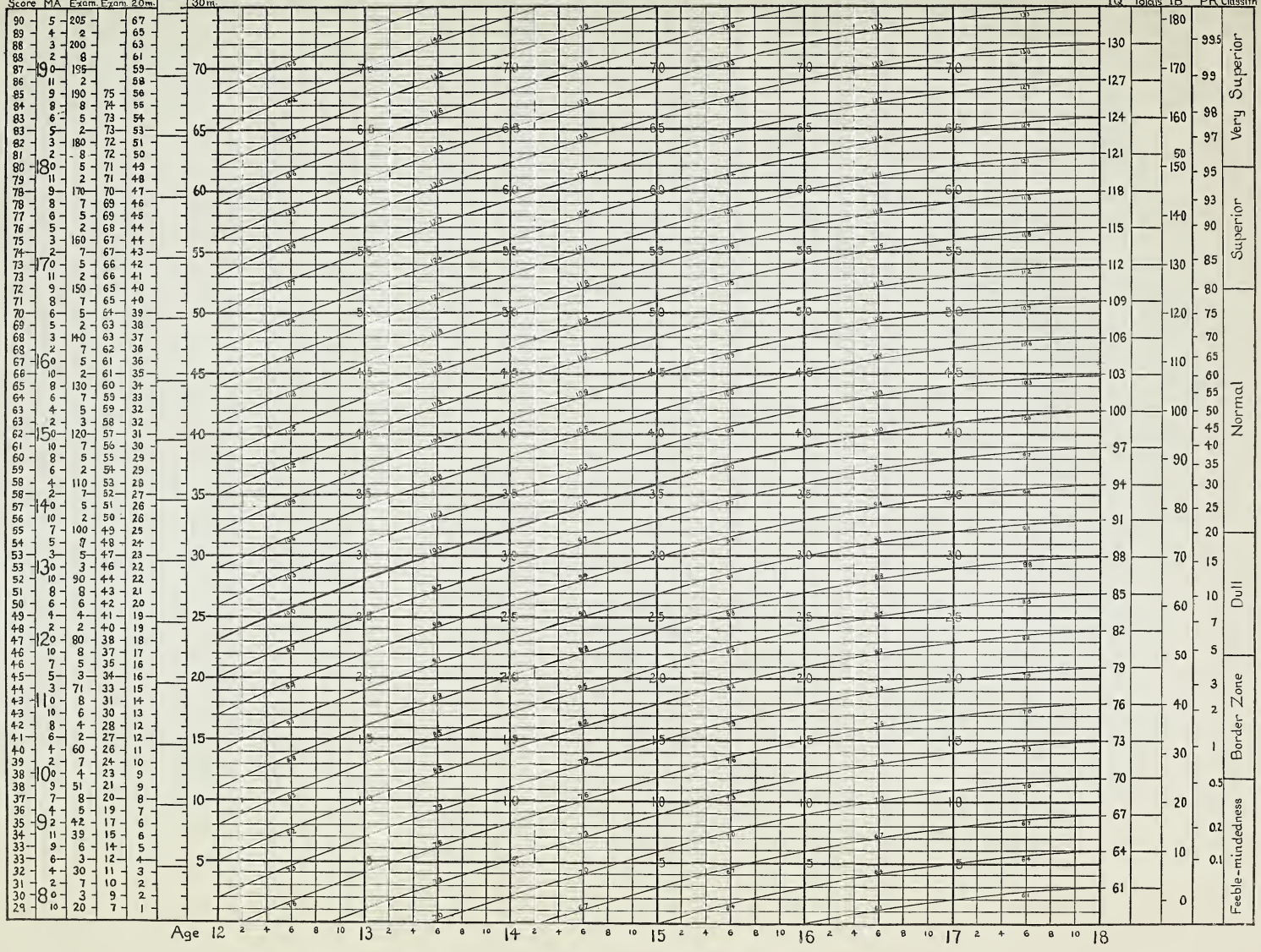
RELATIONSHIP BETWEEN INTELLIGENCE AND VOCABULARY

Earlier the assumption was made that there is probably some relationship existing between inherent or native intelligence of a pupil and the excellence or reverse of his score on the vocabulary test. To discover the extent of this relationship, the scores of as many students as possible on the Otis Self-Administering Mental Test were procured. With some exceptions, this test was given to the students at about the same time as the Vocabulary test. However, the Intelligence Quotients of most of the students from Lamont were obtained from the Alberta School of Education survey of the schools in that town made in March, 1934. About half the I.Q.'s of the Strathcona High School in Edmonton were obtained from the record cards of the pupils in the school files. Care was taken, however, to select only those obtained on the Otis Test. In some cases, these were two or three years old, but as the Intelligence Quotient is considered constant for any given purpose, these were accepted, In Dalemead Rural High School this was written by the students in the autumn of 1934, and these results used.

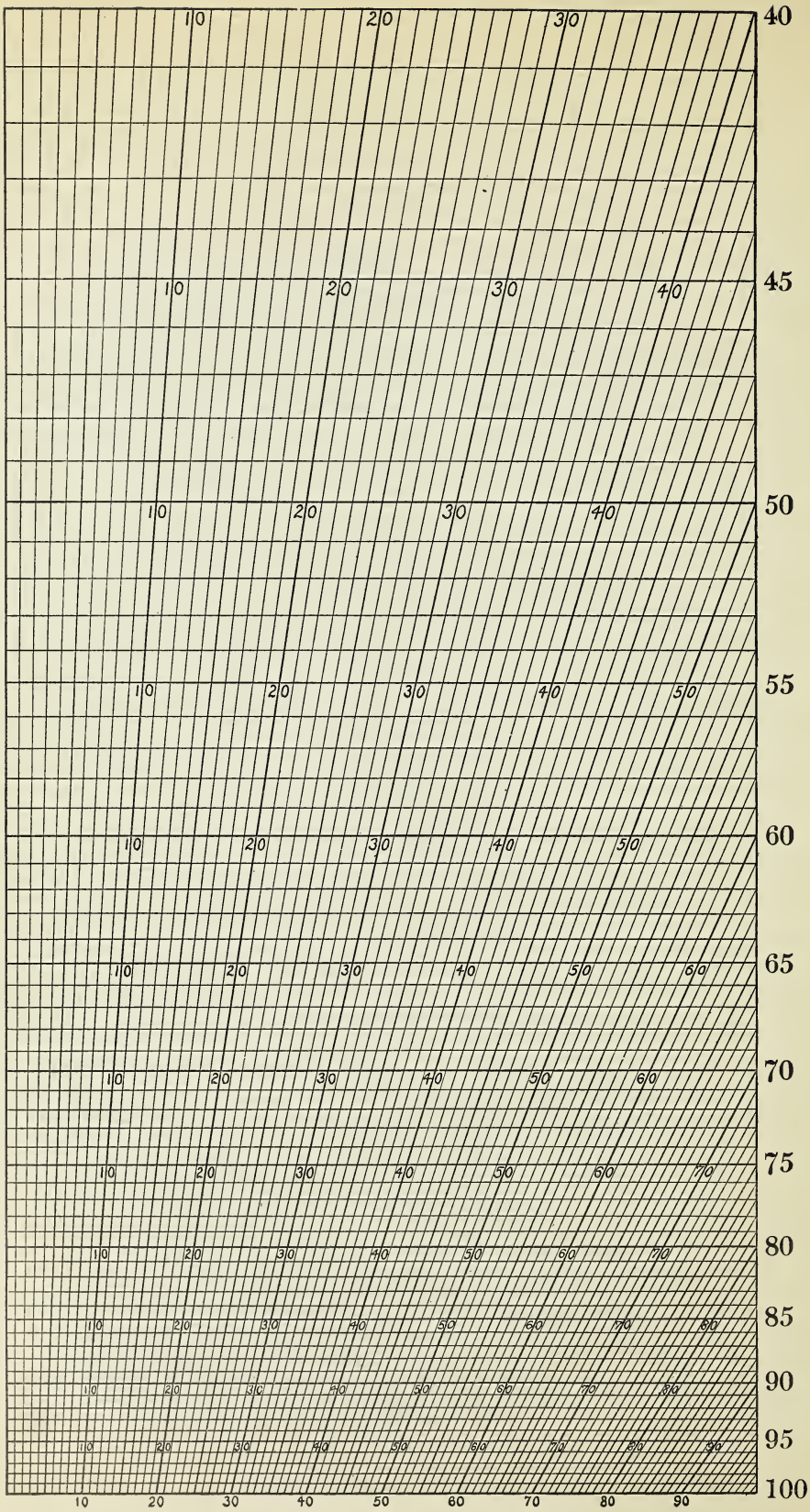
In Grades VII and VIII the Intermediate Examination, Form B, was used, while in the high school grades Form B of the Higher Examination. Although most teachers allowed the students thirty minutes in which to answer the test, a few classes were allowed only twenty minutes. As the test is standardized to both time limits, this is immaterial.

In analyzing the data obtained from these tests, all scores were translated in equivalents of the absolute score on the Higher Examination. For the Intermediate Examination this was done by means of the tables which appear on the Interpretation Chart for Higher Examination; which chart has been included as Table IX. The same procedure was used to convert twenty minute scores into thirty minute scores. Where the I.Q. alone was obtained, as from Lamont and certain cases in Strathcona High School, the pupil's age was considered to be one half year more than that obtained on his last birthday. Then, on the same chart his I.Q. was found and the appropriate score obtained. This would no doubt introduce a few errors by placing some pupils too high and others too low, but these errors would compensate each other.

The reason for using the Higher Examination score or equivalent was that it was desired to have a measure of the student's absolute mental ability without reference to chronological age. On the same chart Binet Mental Age equivalents are given and these might have been used. However, this was not done as the distance from one year to the next is not constant. For instance, the distance from a mental age of nine to that of ten is represented by four seventy-fifths of the total score, whereas the distance from a mental age of sixteen to one of seventeen is represented by eight seventy-fifths of the total score. Vocabulary and intelligence score were obtained from four hundred and ninety-two pupils, the coefficient of correlation between being them computed. Letting $\hat{\theta}$ stand for the score obtained on the Otis Test and V for that obtained on the Vocabulary


Measures of Mental Ability
"T" Binet Adv. Inter. Higher Totals Higher
Score MA Exam. Exam. 20m. 30m.
Measures of Brightness
IQ Totals IB PR Classifn
Very Superior
Superior
Normal
Dull
Border Zone
Feeble-mindedness
Age 12 13 14 15 16 17 18

SCALE CHART



Test, the following measures were derived:

	V	θ
Means:	63.07	40.17
Standard Deviations	16.50	11.83

Coefficient of Correlation $r = .632 \pm .018$

Regression Coefficients θ on V, $m_V = .457$ V on θ, $m_\theta = .882$

Regression Equations:

$V = 63.07 + .882(\theta - 40.17)$ $\theta = 40.17 + .457(V - 63.07)$

In terms of m.a.

$V = 63.07 + .882(M.A. - 15\ 1/36)$ $M.A. = 15\ 1/36 + .457(V - 63.07)$

Correlation ratios: very approximately

θ on V, $Eta_V = .646$ V on θ, $Eta_\theta = .650$

Linearity of Regression:

$(Eta)^2 - r^2 = .017$ $(Eta)^2 - r^2 = .022$

$(Chi)^2 = 4.93$

The coefficient of correlation so obtained is very much higher than that derived between Vocabulary Score and chronological age. Moreover, the Probable Error is relatively small, and the chances are fifty out of one hundred that the correlation lies between r + 1 P.E. and r - 1 P.E.; or between .6 14 and .650. To go still further, the chances are ninety nine out of one hundred that the true r lies between r + 4 P.E. and r - 4 P.E.; or between .560 and .704. In other words, we have a "reasonable" sureness that it lies between the limits indicated (Camp, page 263). This is very much greater than the coefficient of correlation between vocabulary score and chronological age, and indicates that a child's vocabulary depends much more upon his mental maturity than upon the length of time he has been in the world.

The two regression equations give up the vocabulary score a child should attain, given a definite intelligence score; and the intelligence score corresponding to his vocabulary score. These have been translated into Mental Age terms, the value for the mean of θ being taken from the Interpretation Chart mentioned above. It is permissible so to translate these equations as the regression coefficients do not depend on the Otis score, but wholly on the distribution of the frequencies. In this form they are more valuable than in the other, for insofar as the m.a. obtained from this intelligence test corresponds to the m.a. of other tests, these equations will apply to them. The lines which these two equations represent, as well as the column and row means, or intelligence and vocabulary means respectively, have been plotted in Diagram XX.

The measures of linearity of regression derived from the correlation ratios, are materially less than these measures with respect to the computations on vocabulary score and chronological age. This means that the column means cluster more closely around the regression lines, and this is evident from Diagram XX. Here, as in Chapter V, the extremities of the distribution are rather irregular, this being no doubt due to the paucity of cases. In short, proficiency in vocabulary depends much more upon mental than chronological age.

So far we have assumed our populations to be drawn from a universe in which the various measure studied are distributed according to the normal law, and we have normalized our distributions on

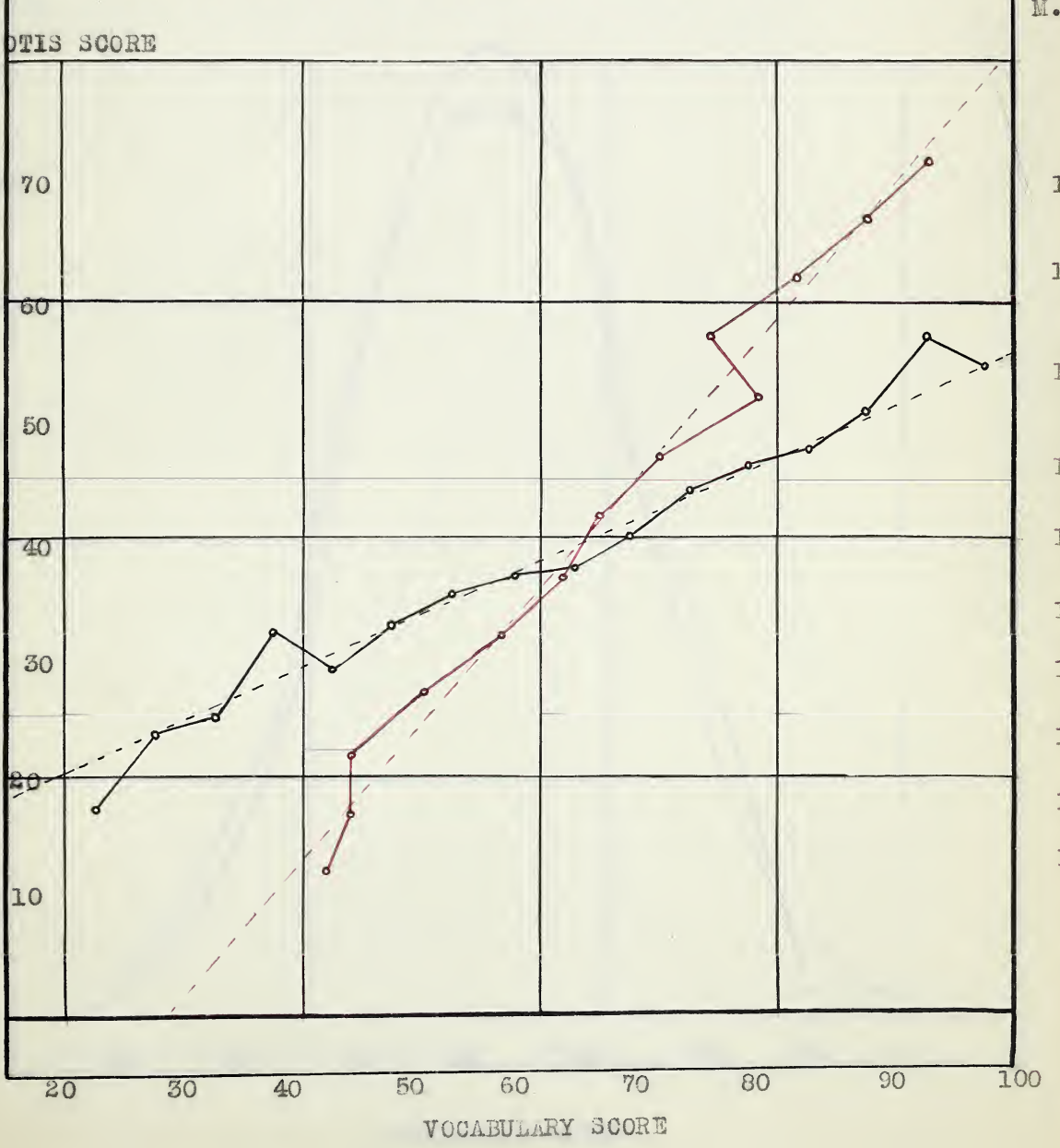
DIAGRAM XX

REGRESSION BETWEEN INTELLIGENCE AND RECOGNIZED VOCABULARY

OTIS TEST MEANS ——— VOCABULARY TEST MEANS ———

REGRESSION OF INTELLIGENCE ON VOCABULARY----O =40.17 -.457(V-63.07)

REGRESSION OF VOCABULARY ON INTELLIGENCE---V =63.07-.882(O-40.17)



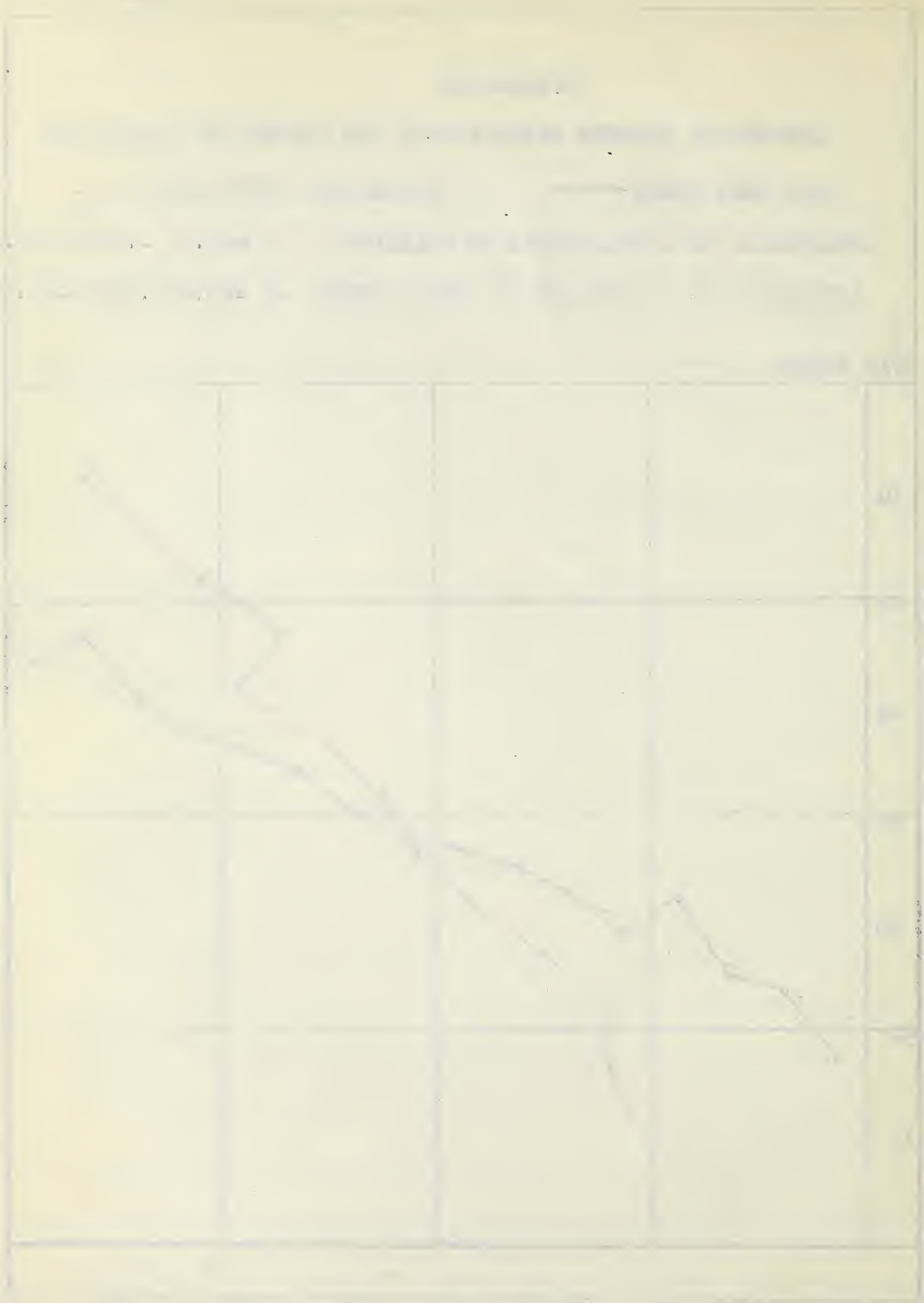
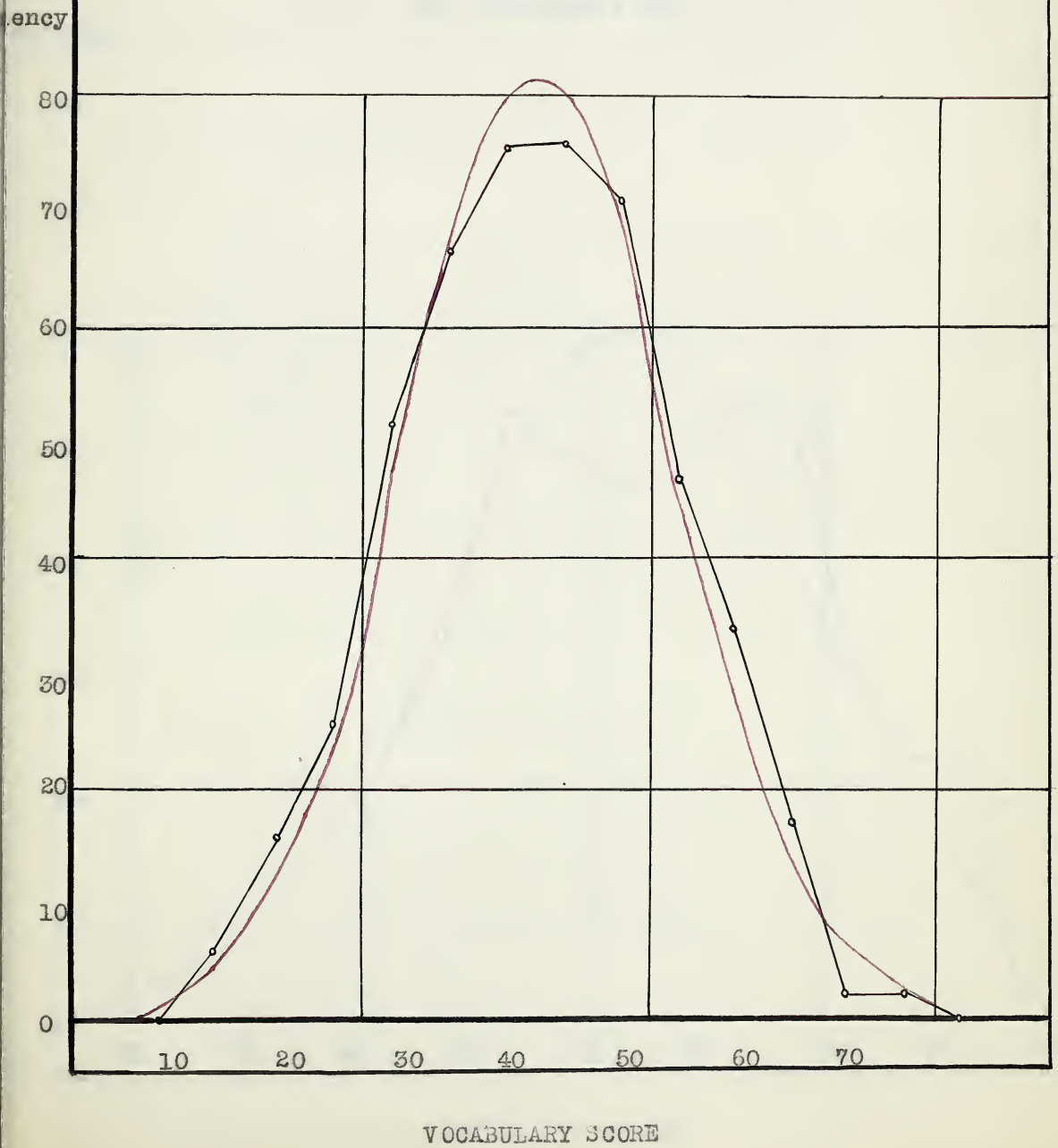


DIAGRAM XXI

DISTRIBUTION OF OTIS SCORES IN GRADES VII TO XII

ACTUAL DISTRIBUTION ——— THEORETICAL DISTRIBUTION ———

NO. OF CASES = 492.



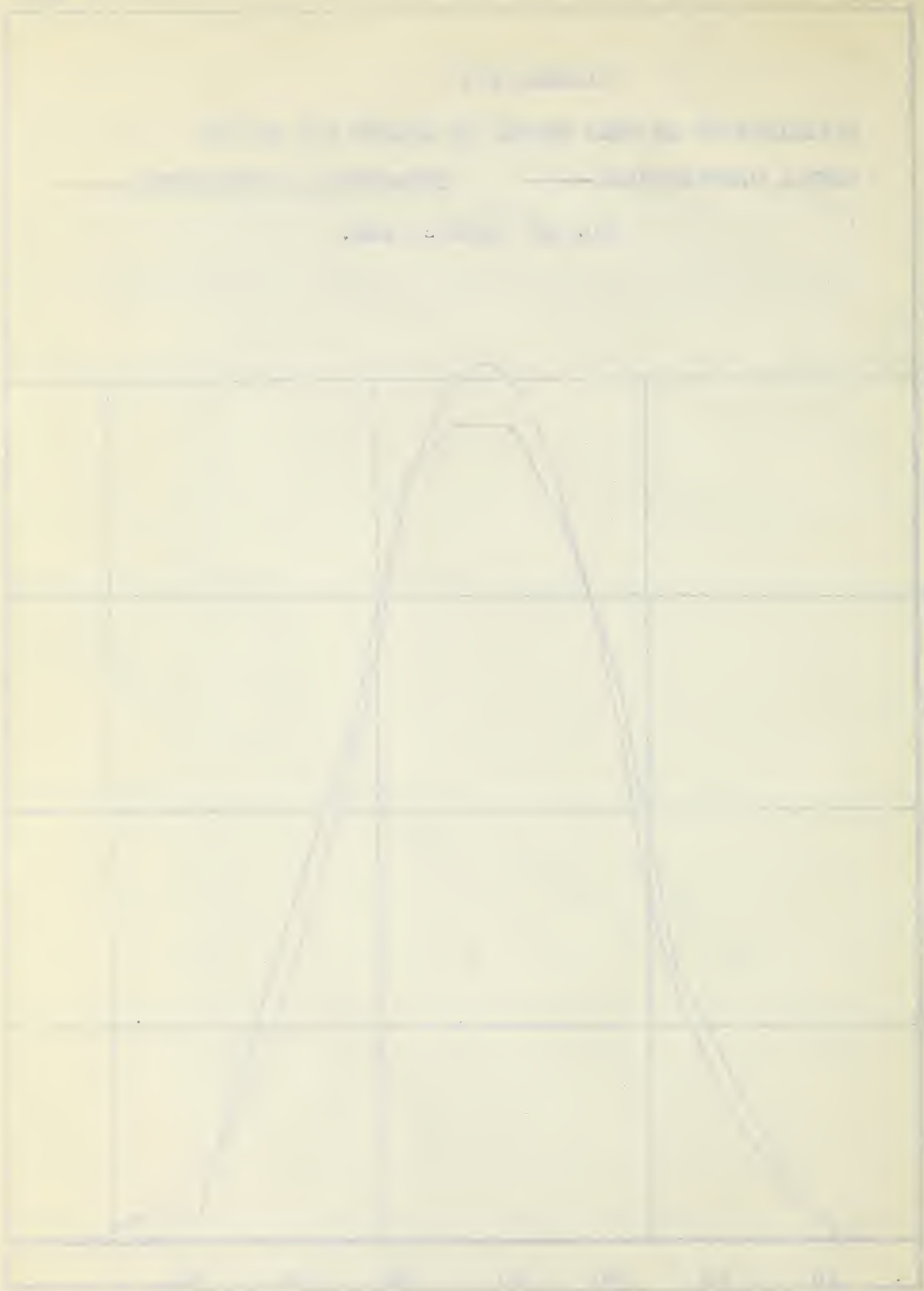
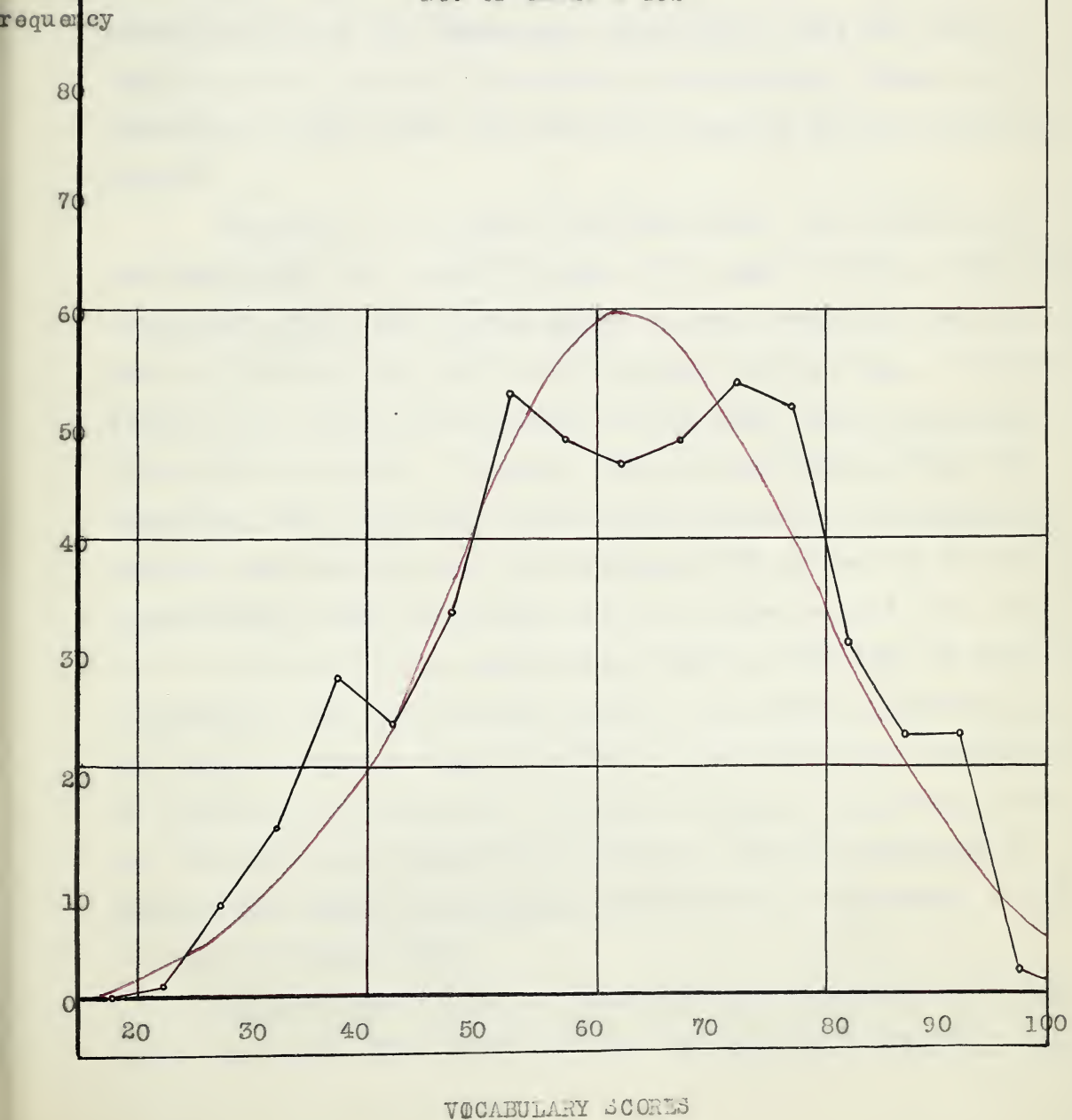


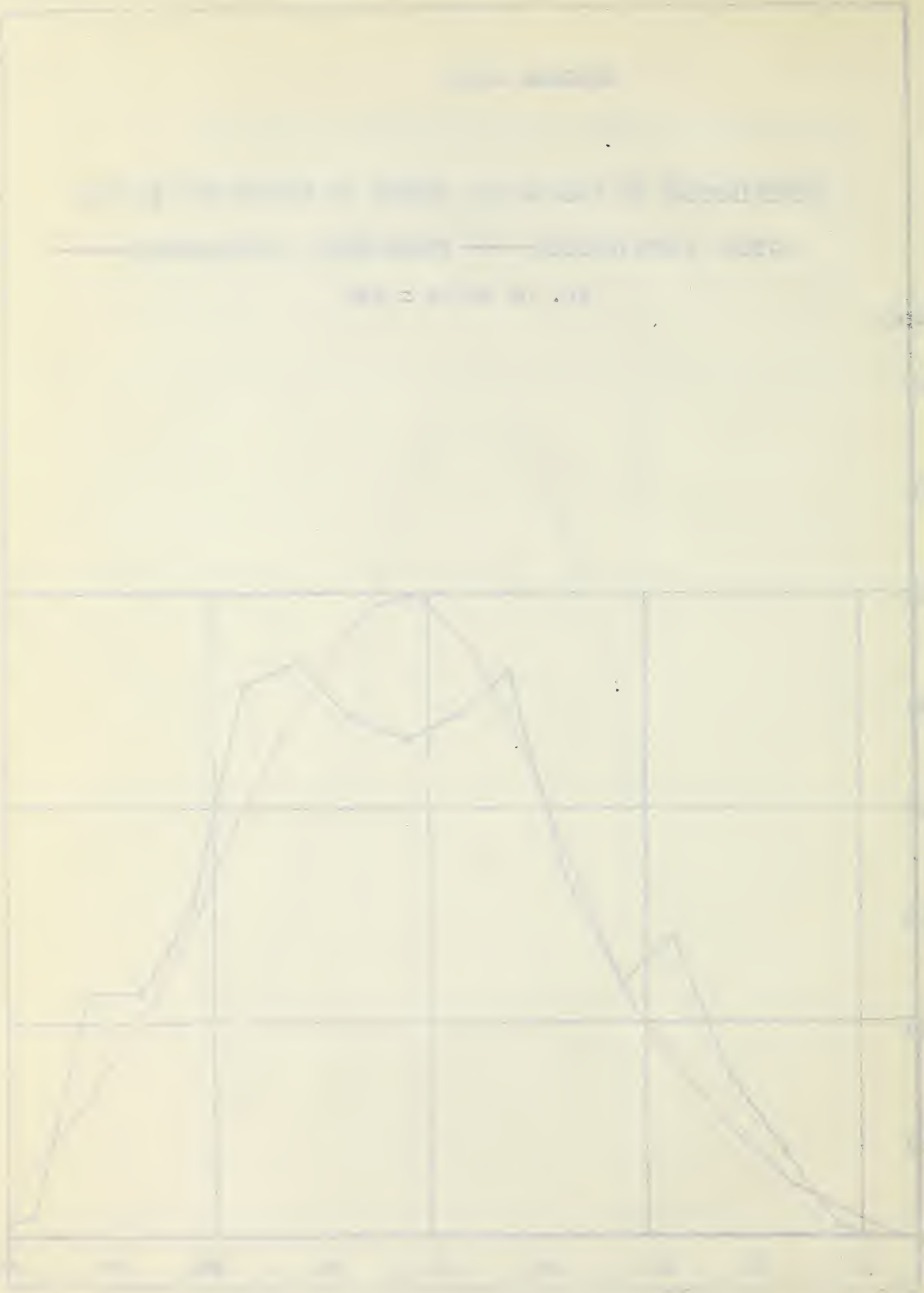
DIAGRAM XXII

DISTRIBUTION OF VOCABULARY SCORES IN GRADES VII TO XII

ACTUAL DISTRIBUTION —— THEORETICAL DISTRIBUTION ——

NO. OF CASES = 492





the basis of that assumption. However, we do know that the distribution of intelligence follows this law very closely and so we are enabled to test our sample by means of the chi test. This is described by Camp in his "Elementary Statistics", page 264 et seq. Speaking of it, he says "It should not be applied, without some modification to cases where the universe is partly inferred from the sample".

Therefore it is quite applicable here. The distribution of the Otis scores was tested by means of it, and it tells, using Camp's condensed tables, that we may expect to get a sample as far from perfect as this one from 91.6 to 98.3 percent of the time; or interpolating in the tables, 95.2 percent of the time. Just how perfect it is we shall see below. However, there is certainly no reason for suspecting that it is not a truly random sample of the population studied; and since it is a random sample with respect to the intelligence test, we may infer that it is to a great extent with respect to the vocabulary test, remembering, however, that here we encounter bilingualism and other factors which do not affect a person's intelligence (although they may affect his manifestation thereof) but do influence his vocabulary. The distribtuion of the Otis scores for the whole four hundred and ninety-two cases is indicated in Diagram XXI, while the analogous distribution of vocabulary scores is shown in Diagram XXII.

As mentioned above, all students were considered to be exactly six months older than they were on their last birthday. The

variates were than grouped according to age, grade and Otis score; and the scores were so arranged that the variates automatically fell into groups with an I.Q. of 100 - 110, 110 - 120, and so on; the I.Q. limits being chosen from Diagram IX with the assumption that all students were six months past their birthday. An illustration will probably make this clearer, so below is given as an example the distribution of all students sixteen years old and in Grade XII.

I. Q.	Otis Score	Frequency
130-140	70.1 -	1
120-130	60.1 - 70.1	2
110-120	50.1 - 60.1	5
100-110	40.1 - 50.1	5
90-100	30.1 - 40.1	1
80- 90	20.1 - 30.1	0
70- 80	10.1 - 20.1	0

The frequencies were then collected by grades and by ages, so that two tables were derived, distribution of I.Q.'s by grades, and distribution of I.Q.'s by ages. These distributions are shown in Tables Xa and Xb respectively. On these tables, "Delta" stands for the coefficient of normality.

Since mental growth after eighteen and indeed after an earlier age in the opinion of some psychologists is very gradual, all frequencies above seventeen years have been grouped together.

At the lower end of the scale the group limits for the age of eleven years, six months, had to be extrapolated, but this seemed reasonably accurate, using Table Ix.

In all, four hundred and ninety-nine intelligence scores were tabulated; this being somewhat more than in the correlation. Seven students in the classes tested wrote the intelligence test but not the vocabulary test.

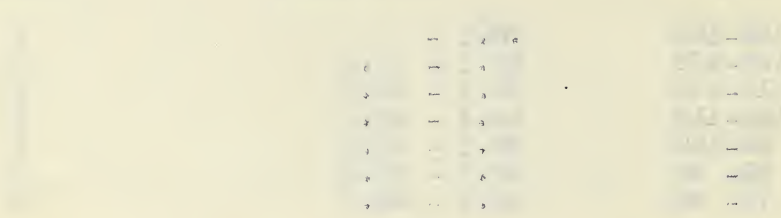


TABLE X (a)

DISTRIBUTION OF INTELLIGENCE QUOTIENTS BY GRADES

I.Q.	XII	XI	X	IX	VIII	VII	TOTALS
130-140	1			2	1	2	6
120-130	4	13	1	10	12	3	43
110-120	23	21	9	18	19	19	109
100-110	30	26	25	31	22	16	150
90-100	16	22	27	26	15	13	119
80- 90	3	13	11	15	9	6	57
70- 80		3	2	5	4	1	15
TOTALS	77	98	75	107	82	60	499
DELTA	.51	.23	-.07	.14	.29	.30	.24

TABLE X (b)

DISTRIBUTION OF INTELLIGENCE QUOTIENTS BY AGES

I.Q.	over $17\frac{1}{2}$	$17\frac{1}{2}$	$16\frac{1}{2}$	$15\frac{1}{2}$	$14\frac{1}{2}$	$13\frac{1}{2}$	$12\frac{1}{2}$	$11\frac{1}{2}$	
130-140			1			2	2	1	
120-130	2	6	5	7	10	10	3		
110-120	13	17	11	16	22	17	9	4	
100-110	19	19	24	31	25	21	11		
90-100	17	15	33	23	19	11	1		
80- 90	12	7	15	11	9	3			
70- 80	1	1	7	5	1				
TOTALS	64	65	96	93	86	64	26	5	499
DELTA	.06	.29	-.16	.15	.33	.56	.62	1.00	.24

Sandiford tells us (Educational Psychology, page 162) that in a normal population, thirty percent of the cases have an I.Q. of 100 to 110, thirteen percent from 110 to 120, six percent from 120 to 130, one percent above 130; and corresponding percentages for I.Q.' s below one hundred. Upon these facts has been founded our method of deriving of coefficients of normality for our various ages and grades In chapters six, and seven, the coefficient of normality was derived from the relative frequencies measured in terms of sigmas above and below the mean. However, our coefficient is a versatile animal, and is equally applicable here. Using the same system as previously, but measuring our relative frequencies from an I.Q. of 100, they were compared with the relative frequencies of I.Q.'s as stated in Sandiford. Since our coefficient is really a measure of difference between our sample and the theoretical universe from which it was drawn, it has for the time being been called delta.

Our reason for tabulating our variates according to the O tis score rather than computing the I.Q. is that the limits to the classes fall between two even numbers. In the example shown above they fall at 70.1, 60.1, 50.1, etc. Since the Otis score is in even numbers, by this means we avoid having any of the cases falling at exactly at an I.Q. of 140, 130, 120, etc., corresponding to the Otis scores just mentioned, and so there are no partial frequencies in any of the groups.

Table X (a) shows us that for every grade except Grade X the coefficient of normality is positive, meaning that the intelligence of the students in the grades is higher than that of the

whole population from which our universe is drawn. In Grade XII it is quite high, being .60. As this is indubitably the heaviest school year, it is natural that a higher than average intelligence is necessary to cope with it. In fact, the whole high school seems to be a continuously more selective factor in weeding out the duller students, who begin to drop out of school once they reach Grade VII or VIII. This does not account for the high coefficients in Grades VII and VIII or at least only partially. It was surmised in Chapter V that whatever academic acceleration takes place occurs below these grades, so that we have here a certain number of young and clever students. Table X (b) reveals this to be the case, as the delta from elven and one half to thirteen and one half years run from 1.00 to .56 respectively. At the eleven and one half year level, however, there are only five students, which is too small a number to permit generalizations.

In Diagram XIX it was shown that the age means on the vocabulary Test were in the lower grades studied considerably above the corresponding grade means, and it was suspected that the younger pupils were accelerated and therefore influencing grade means above those corresponding to their own ages. Now we know that the younger pupils are considerably above the average in Intelligence Quotients, and this is no doubt what has happened.

It is rather surprising that the deltas both for Grade X and for the corresponding age (16½) are below the deltas for the grades immediately above and below. We find what is presumably a result of this in Diagram XVIII, where there is a sudden drop in the third

quartile for Grade IX, and this is reflected in the semi-inter-quartile range for this grade.

In discussing the age means of the vocabulary scores, the hypothesis was made that these means were unduly lowered by the less brilliant students being retarded in the higher grades while their more brilliant contemporaries have passed on to other fields of endeavor. Had the coefficient of normality of the intelligence quotients for these years been negative, this hypothesis would have been directly corroborated. Since, however, for both seventeen and one-half years, and for more than seventeen and one-half years it is positive, it would seem that our hypothesis is fallacious. Such, however, is not necessarily so, for while those still in Grade XII are bright, those of the same age who have already graduated from this grade may be even more so. This is indicated by two facts. First, the delta of intelligence quotients drops from .29 to .06 in progressing from an age of seventeen and one-Half to higher ages, indicating that the more clever of these older students have already left school. Second, the delta for Grade XII, .60, is higher than for either of these age groups, indicating the presence of younger brighter students advancing rapidly through their academic career.

But if the grade norms for the Vocabulary Test, especially at the higher levels are too low, the age norms are probably too high. They are applicable no doubt to students in the grades studied but since students in these grades are somewhat more intelligent than the average of mankind, these age norms should not be applied either to students below Grade VII, or to children who have left

school before completing a reasonable amount of high school work.

In discussing the low scores obtained by bilingual students, it was suggested that, since the Otis Test is entirely a reading test, should these students obtain low scores on the Otis Test, that might be due to a deficiency in English rather than a meagre accomplishment in English being due to a low intelligence.

The Otis scores of eighty-two bilingual students were studied by dividing them into I.Q.'s as was done with all the scores above, and it was found that on the average bilingual students do make lower than normal scores on the Otis Test. The Central European group, comprising forty-seven cases is the only one large enough to consider separately. For this group the Coefficient of Normality on the Otis Test is -.26, compared to -.35 on the Vocabulary Test. For the whole bilingual population, the delta for the Otis Test was -.07, compared to -.32 on the Vocabulary Test. Although the two groups are not wholly comparable, the Otis group comprising only about seventy-five percent of Vocabulary group; admitting that the Otis test is a true measure of the mentality of bilingual students, it would seem that the bilingual children are hampered on the Vocabulary Test beyond the amount which would be expected from their lower intelligence. This is so because both for the Central European group and especially for the whole foreign-speaking group, the deltas for the Vocabulary Test are appreciably lower than for Otis Test.

CHAPTER IX

SEX DIFFERENCES IN VOCABULARY.

It is not pretended that this section is an exhaustive and complete inquiry into the problem of vocabulary differences between the two sexes. It was not expected that any great difference would be found, and accordingly only a superficial examination of the problem has been attempted.

The method used was that of matched pairs. In matching a boy with a girl, rather rigorous requirements had to be met. Both had to have passed the same birthday, i.e., they had to be within a year of each other in chronoligcal age. They were chosen from the same school, the same grade, and the same occupational group. Students speaking English only were considered. Finally, the greatest difference in the Otis score allowed between each member of the pair was three, representing a difference in the mental age of six to eight months. The various pairs are represented below, with the respective scores. To obtain the differences, the girl's score was subtracted from the boy's.

School	Grade	Age	Vocab. Boys	Vocab. Girls	Diff.	Otis Boys	Otis Girls	Diff.	Occupational Group
Western Canada	X	16	70	61	9	33	33	0	Professional
"	"	16	62	71	-9	37	40	-3	Commercial
"	"	16	77	52	25	41	39	2	Managerial
Bellevue	IX	15	66	55	11	36	34	2	Unskilled
"	VIII	14	69	73	-4	48	48	0	"
Delemead	XI	16	91	79	12	58	58	0	Agricultural
Garneau	IX	14	73	75	-2	53	56	-3	Clerical

CHAPTER X

VOCABULARY SCORES AND OTHER ACHIEVEMENTS IN ENGLISH.

One would naturally expect that a person who did very well on the Vocabulary Test would also excell in other branches of English. An effort was made to discover if this was so, and to what extent.

The scores obtained by the students in Grades VII, VIII, and IX on the New Stanford Reading Test were studied. Although this test is designed primarily for Grades IV to IX, it has age norms up to the age of nineteen years, two months, and it was the original intention to give this test to all the students. However, it was soon seen that from Grade X up most students obtained almost perfect scores on this test, and as it did not measure any differences in ability in the higher grades, the test was administered to only some of the students in these grades, and even these scores have been discarded.

As this test was administered during February, the norm would be for Grade 7.5, 8.5 or 9.5 as the case might be. The means and medians of all the students tested in these grades were found and the results were as follows:

	VII	VIII	IX
No. of cases	95	54	91
Given Norm	86	93	99
Mean	87.45	94.26	103.54
Sigma	10.9	5.82	9.6
Median	87.5	9.572	105.83

Taking either the mean or the median, as the average score, the norm seems satisfactory for Grade VII students, a trifle low for Grade VIII and very much too low for Grade IX. The norms given stop at Grade 10.0, where the norm is 102, so we see that according to the New Stanford Reading Test, the Grade IX students are accelerated to the extent of one whole grade. In the higher grades the discrepancy was even worse, scores frequently being above the entire range of the norms.

There may be several reasons for these discrepancies. Canadian students may be better trained in reading than their American contemporaries. Methods of teaching reading may have improved to such an extent in the last few years as to make the norms obsolete. This is at least plausible in view of the great strides forward that have been made in the scientific study of this subject. And because the given norms above Grade 8.4 and a score of 92 have been extrapolated, error in computing them might easily enter.

The New Stanford Reading Test is divided into two sections, and so arranged that equivalent scores on each test represent equivalent standards of achievement. In determining the averages above the arithmetical mean of the two scores was used. The first part of the test is on paragraph meaning. The subjects read paragraphs that became increasingly difficult, and in the last sentence, there are one or more blanks to fill in, this testing comprehension of the matter read. There are eighty such blanks and twenty five minutes is allowed to complete this section of the test. Although

School	Grade	Age	Vocab. Boys	Vocab. Girls	Diff.	Otis Boys	Otis Girls	Diff.	Occupational Group
Strathcona	IX	14	73	75	-2	53	56	-3	Managerial
Tofield	IX	19	73	71	2	47	44	3	Agricultural
Turin	VII	13	43	46	-3	30	27	3	Agricultural
Victoria	XI	16	68	69	-1	43	46	-3	Commercial
"	"	17	89	76	13	53	52	1	Commercial
TOTALS					51			-1	
AVERAGES					4¼			-1/12	

Although in half the cases the girls are superior to the boys, yet the highest vocabulary difference in favour of a girl (9) is equalled by one boy and exceeded by four, only one having a difference of less than nine over his mate where he is superior at all. Moreover, the average difference in mental ages is scarcely equivalent to one week against the girls, whereas the average difference in vocabulary score is equivalent to about one year in favour of the boys. Although these differences may be due to chance, considering the fewness of cases, considering the size of the vocabulary average difference, it seems unlikely. We cannot say that more boys than girls do well on this test, but when the boys do exceed the girls, they do so by a substantial margin. Perhaps this is due to the boys having a wider range of interest than the girls, such as in sport, mechanics, adventure, etc. However, a more thorough investigation is necessary before our information can be considered wholly reliable.

this is theoretically a purely objective test, in practice it is not perfectly so, as it is often difficult when a child gives an unorthodox response to decide whether he has understood what he has read or not.

The second part of the test, for which ten minutes are allowed, is a word test. It consists of eighty very short sentences in which the last word must be chosen by underlining from five choices presented, and so is very similar to our Vocabulary Test, but much less difficult.

Two hundred and twenty-seven students in the Grades VII, VIII, and IX answered both the Reading and the Vocabulary Tests, giving us two scores on the former and one on the latter. Each of these scores was correlated with the other two. The highest coefficient of correlation obtained was between the Vocabulary Test and the Word Section of the Reading Test, being .85 ± .012. It is not surprising that the greatest concomitance should be found between these two tests, as they both measure the same thing; but it is very gratifying that the correlation should be so high with such a relatively low probable error, as this gives us evidence as to the reliability of our test beyond that furnished by the Thorndike Test of Word Knowledge.

The next highest coefficient of correlation was that between the two sections of the Reading Test, being .81 ± .016. The two sections of the test measure abilities that are somewhat different yet with many elements in common, and accordingly this coefficient is indicative as to the validity of the two tests.

The lowest coefficient although still high, being .78 ± .017 was that between the Vocabulary Test and the paragraph section of the Reading Test. Comparing this with the coefficient between the two sections of the reading test, we may safely say that the Vocabulary Test is almost if not quite equal in validity to the word section of the Reading Test. Indeed, considering the Probable Errors, it may easily happen that further study would prove the Vocabulary Test to be fully as valid as the Reading Test.

It is interesting to note that the students tested did somewhat better and with less deviation on the paragraph than on the word section of the Reading Test. The respective means for the two sections were 96.02 and 93.07; the respective sigmas 12.8 and 14.00.

The coefficient of correlation between the Vocabulary Score and the Hudelson composition score was also obtained. In this case only those students who wrote the Vocabulary Test in February and the composition in October were considered, but as there were four hundred and twenty-two of these, the sample is amply large. The coefficient of correlation here was found to be .66 ± .019. That this is lower than those obtained with respect to the Reading Test is not surprising. In the first place, the subjectivity of scoring the compositions makes complete accuracy impossible. The very fact that the Hudelson score is obtained by a scale rather than a test might raise doubts as to the soundness of correlating the Hudelson and the Vocabulary scores by the product moment method of correlation. However, the Hudelson rating is supposed to measure

composition merit quantitatively, so this procedure was thought permissible.

Another reason why we should not expect a high correlation between the Hudelson rating and the Vocabulary score is that composition merit depends on many things besides extent of vocabulary. One with a meagre vocabulary may work carefully with close attention to spelling and other mechanical details, and so achieve a high score. On the other hand, a clever student, bored with the whole proceeding, may write carelessly, with but little consideration to punctuation and grammar.

Considering these things, the coefficient obtained is really quite high and speaks well for the Hudelson scale, despite its imperfections.

The scores obtained on the Vocabulary Test were also compared with the written vocabulary as shown by the themes themselves. The more words a person uses that are not in the most common five hundred, the richer in general we may assume is his vocabulary of expression. Accordingly the number of words chosen from the most common five hundred is a rough inverse measure of the extent of the expressive vocabulary. This number was accordingly correlated with the vocabulary score. The same compositions were used here as were used in finding the coefficient of correlation between the Hudelson score and the Vocabulary test, so that with over four hundred cases the sample may be considered adequate. The coefficient was found to be -.22 ± .031. There is apparently only a very slight connection between the words a child knows and those he uses. This is what we should expect, in view of the fact that

growth in written vocabulary, as shown in Diagram VIII, IX, and X, is very slow indeed.

CHAPTER XI

CONCLUSION

The results of this summary may be recapitalized very briefly. By the time a child has reached Grade VII, his vocabulary that he uses in every day life is practically complete, and there is henceforth very little growth and transfer of words from recognition only to use. Furthermore, about eighty percent of the words he uses come from the most common five hundred in the language. This may go up to as high as ninety-five or as low as sixty percent, but such cases are exceptional. Perhaps this is as it should be, and certainly one is in no dangerof being misunderstood when one sticks to such common words. If this is a desirable state of affairs, however, much time spent in learning the spelling and inflections of the more uncommon words might conceivably be used to better advantage in practising using such words.

The growth in recognized vocabulary is much more encouraging. As on the average one word in our test represents two hundred in the Word Book, in the grades studied growth in recognized vocabulary takes place at the rate of about one thousand words a year or more. However, there are wide variations within each grade; only about twenty percent of the students being at their own grade level, with about forty percent below and forty percent above the grade level.

These variations are due to many factors, perhaps the chief of which is the inherent intelligence of the pupils. Chronological age plays some slight part, and perhaps sex also. Environment, as

as revealed by the parent's occupation is quite important, the children coming from professional homes having a great advantage over those from farm homes, for instance. The fact that bilingual children are definitely weaker in their accomplishment in the Vocabulary Test indicates that familiarity with the language plays an important part in the acquisition of new words, as would normally be expected. Therefore, the teacher of foreign-speaking students should make extra efforts to compensate for this handicap by frequent practice in written and oral composition and by as much time devoted to supplementary reading as possible. Proficiency in the use of the language would seem to be more important than mechanical excellence, desirable as that may be.

Excellence in recognized vocabulary is accompanied by excellence in other fields of English, notably reading. This is no doubt true as these two skills are greatly interlinked and largely dependent upon one another. And in general those who know the most words will express themselves better than those who do not, but they will not use a greater variety of words to do so, or at least only to a small extent.

The vocabulary test designed for this work seems to be a good measuring tool and well adapted to the high school grades. It seems to be well graded in difficulty to show differences in ability and is difficult for even the most proficient high school student. This is shown by the fact that out of over five hundred cases, not one student made a perfect score, and only two reached ninety-five or

over. It possesses a high degree of validity, and would seem to be very reliable. The test is easily understood and can be answered quickly. The scoring can also be done very rapidly; much more so than for any of the other tests used in this study.

Tentative grade norms have been established, and also age norms, but the latter appear to be too high at the lower levels studied and too low at the upper levels. These could be adjusted at the bottom by obtaining scores from children of normal ability in the grades below Grade VII, and corrected for the top possibly by using scores obtained by University and Normal school students, and by students in technical institutes and agricul tural schools.

Many courses are open for investigation in this field. In the first place, it would seem necessary to obtain norms for the American tests used that are locally applicable, and probably for other tests of similar nature. The Hudelson norms in the lower ranges and the New Stanford Reading norms in the upper ones seem especially unsuitable.

There also seems to be a place for a reading comprehension test similar to the New Stanford, but much more difficult. This might be obtained by using technical material and material dealing with abstractions, although probably in such cases, native intelligence would become increasingly more important.

A thorough study of the handicap suffered by bilingual students and remedial measures to overcome them would be a valuable contribution to educational science, especially in a country like

this where so many students come from foreign-speaking homes. In such an investigation, French students should especially be studied, if only for the reason that French is a Canadian language as much as English. It would also be interesting to know how French-speaking bilingual students compare, with respect to their command of French, with those who speak and use that language only.

- - - -

BIBLIOGRAPHY

1. Thorndike : "Teacher's Word Book of 20000 Words"

2. MacLeod : "A Statistical Study of the Validity of Objective and Essay Examinations in History II"

3. Brandenburg : "Psychological Aspects of Language". Journal of Educational Psychology vol. 9, p. 313.

4. OBrien : "Vocabulary of High School Students in Written Composition" Journal of Educational Research , vol 11, page. 344.

5. Witty and Fry : "Vocabulary Content of Compositions Written by College Students" Journal of Educational Research, vol. 19, page 135.

6. Symonds and Lee : "Studies in the Learning of English Expression" Teacher's College Record, vol. 31, page 50.

7. Sims : "Reliability and Validity of Four Types of Vocabulary Tests". Journal of Educational Research. Vol. 21, page 94.

8. Caldwell and Moury: "The Essay vs. the Objective Examination as Measures of the Achievement of Bi-lingual children". Journal of Educational Psychology. Vol 24, page 696.

9. Fritz and Rankin: "The English Handicap of Junior High-School Pupils from Foreign Speaking Homes and Remedial Measures". Journal of Educational Research. Vol 27.

10. Camp : Elementary Statistics.

11. Sandiford : Educational Psychology.

B29743